Map-based cloning of the gene *albostrians* in barley (*Hordeum vulgare* L.)

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biologie

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
M. Sc. Mingjiu Li

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Lebenwissenschaftliche Fakultät
Prof. Dr. Richard Lucius

Gutachter: 1. Prof. Dr. Thomas Börner
2. Prof. Dr. Christian Schmitz-Linneweber
3. Prof. Dr. Andreas Graner

Tag der mündlichen Prüfung: 29.07.2015

Table of Contents

List of Figures

List of Tables

Summary

Variegated mutants provide excellent tools to study the mechanisms of chloroplast biogenesis and the nature of communication among the three DNA-containing compartments of a plant cell, i.e. nucleus, chloroplast and mitochondria. The *albostrians* mutant, as one of the most prominent examples of variegated mutants in plants, was extensively characterized over the last three decades for a broad range of aspects, including morphology of the mutant phenotype, gene expression within the nucleus, chloroplast as well as mitochondria. All previous research was performed lacking the information about the underlying functional gene. Thus the focus of this study was to isolate the gene *albostrians* by forward genetics and functionally characterize the gene.

Map-based cloning was employed for identification of the *albostrians* gene. Starting with mapping in two small F_2 mapping populations, MM4205 and BM4205, the locus could be assigned to the long arm of barley chromosome 7H. High-resolution genetic mapping in conjunction with extensive marker saturation allowed to reduce the genetic target interval iteratively from initially 14.29 cM to finally 0.06 cM by analyzing a total of 1344 F_2 plants. A single gene could be identified in a physical distance of 46 Kbp between the closest flanking genetic markers. Functional annotation of the deduced protein revealed it to represent a member of the *CMF* gene family of putative transcriptional regulators comprising DNA binding or protein-protein interaction properties. The identified candidate gene was first confirmed by screening an EMS-induced TILLING population derived from barley cv. ‘Barke’. Among the 42 identified induced mutations a single mutation introduced a premature stop codon potentially resulting in a shorter protein upon translation. Progeny of this heterozygous mutant segregated for green and *albino* plants. The *albino* phenotype was perfectly linked with the homozygous state of the stop codon mutation in 245 M_4 offspring of five heterozygous M_3 plants of the mutant family. Transient transformation by biolistic co-bombardment of barley epidermal cells with an ALBOSTRIANS::GFP fusion protein and an mCherry labelled organelle marker pt-rk-CD3-999 revealed the ALBOSTRIANS protein is targeting to plastids and nucleus. The result indicated that the gene *ALBOSTRIANS* might have its function in regulating gene expression in the nucleus as well as the plastid. Hence, further

characterization of the ALBOSTRIANS protein can be expected to provide a better understanding of the mechanisms underlying chloroplast biogenesis.

Key words: *albostrians* gene, forward genetics, genetic mapping, map-based cloning, TILLING analysis, chloroplast biogenesis

Zusammenfassung

Variegierte Pflanzen sind exzellente Objekte für Studien zur Chloroplastenbiogenese und zur Natur der Kommunikation zwischen den drei DNA-enthaltenden Kompartimenten der Pflanzenzelle, Zellkern, Chloroplast und Mitochondrion. Die Gerstenmutante *albostrians*, eines der prominenten Beispiele variegierter Pflanzenmutanten, ist in den letzten drei Dekaden extensiv unter sehr verschiedenen Aspekten, wie Morphologie des Mutantenphänotyps und der Genexpression im Zellkern, den Chloroplasten und Mitochondrien, charakterisiert worden. Diese früheren Untersuchungen wurden ohne Informationen zum verursachenden Gen durchgeführt. Deshalb stand die Identifizierung des *albostrians* Gens durch forward genetics und seine funktionelle Charakterisierung im Fokus der vorliegenden Arbeit.

Die Identifizierung des *albostrians* Gens erfolgte mittels Karten-basiertem Klonieren. Begonnen wurde mit der Kartierung in zwei kleinen F_2-Kartierungspopulationen, MM4205 und BM4205, die zur Lokalisierung des Genes auf dem langen Arm von Gerstenchromosom 7H führte. Durch Kartierung mit hoher Auflösung in Verbindung mit extensiver Markersättigung konnte der betreffende DNA-Bereich schrittweise von anfangs 14,29 cM auf schließlich 0,06 cM eingeschränkt werden, wobei insgesamt 1344 F_2-Pflanzen analysiert wurden. Zwischen den nächsten flankierenden genetischen Markern konnte in einem Bereich von 46 Kbp ein einzelnes Gen identifiziert werden. Durch Sequenzvergleich des abgeleiteten Genprodukts mit Einträgen in Datenbanken konnte das Protein der *CMF*-Genfamilie putativer Transkritionsregulatoren mit DNA-bindenden oder Protein-Protein-Wechselwirkungs-Eigenschaften zugeordnet werden. Eine erste Bestätigung der Identität des Kandidatengens mit dem *albostrians*-Gen konnte durch Analyse einer EMS-induzierten TILLING-Population (abgeleitet von der Gerstensorte ‚Barke') erreicht werden. Unter den 42 gefundenen induzierten Mutationen gab es eine Mutation, die zu einem vorzeitigen Stopcodon und damit nach der Translation potenziell zu einem verkürztem Protein führt. Die Nachkommenschaft dieser heterozygoten Mutante spaltete in grüne und albino Pflanzen auf. Der albino-Phänotyp war perfekt mit dem homozygoten Status der *nonsense*-Mutation in den untersuchten 245 M_4-Nachkommen von fünf heterozygoten M_3-Pflanzen der Mutantenfamilie verbunden. Nach transienter Transformation von Gerstenblatt-Epidermiszellen mittels biolistischem Cobombardement von ALBOSTRIANS::GFP-Fusionsprotein mit dem

mCherry-markierten Organellenmarker pt-rk-CD3-999 konnte die Lokalisation des ALBOSTRIANS-Proteins in den Plastiden und im Kern beobachtet werden. Das *albostrians*-Gen könnte also für die Genexpression sowohl im Kern als auch in den Plastiden funktionelle Bedeutung haben. Weitere Untersuchungen des *albostrians*-Gens und seines Produkts lassen bessere Einsichten in die Mechansimen der Chloroplastenbiogenese erwarten.

Schlagwörter: *albostrians* gen, Vorwärtsgenetik, genetische Kartierung, Karten-basiertem Klonieren, TILLING analyse, Chloroplasten biogenese

1 Introduction

1.1 Plastid, the chloroplast and photosynthesis

On earth, plant and algae are photoautotrophic organisms that have the ability to fix light energy and convert it into chemical energy. This process called photosynthesis, is a fundamental basis of energy's flow through life, food chains and energy pyramids. Photosynthesis occurs entirely within the chloroplast. As the most noticeable organelle in plants and green algae cells, it serves as an 'energy factory' for supporting the activities of most living organisms on earth.

Plant chloroplasts originated from a cyanobacterial ancestor, through an endosymbiotic process (Raven and Allen, 2003; Yagi and Shiina, 2014). Chloroplasts belong to a family of plant organelles called the plastid, which represents a variety of inter-convertible forms depending on the differentiation of the respective cell type (Figure 1-1). Most of the distinct plastids are derived from undifferentiated proplastids which are found in meristematic and undifferentiated cells. Each meristematic cell contains around 10 to 20 of such proplastids (Pyke and Leech, 1992). Proplastids are colorless and vary in size between 0.2 and 1 µm. When seedlings are grown without any light, proplastids differentiate into etioplasts, containing an undeveloped internal membrane system with semicrystalline structures called the prolamellar body. Upon illumination, the etioplast can develop into a functional chloroplast carrying stacks of membranes – the thylakoids, where most of the photosynthetic protein complexes are situated. In comparison to chloroplasts, all other plastids are non-photosynthetic while still representing important sites for the biosynthesis of metabolites like starch, fatty acids and amino acids in a broad range of plant tissues (Neuhaus and Emes, 2000). One such kind of plastid, the amyloplasts, can be found in root cells where it serves as a storage compartment of starch granules. Besides roots, cells of the storage tissues such as endosperm, tubers and cotyledons also contain amyloplasts. In addition to amyloplasts, elaioplasts represent another form of colorless plastids and are specialized in storing lipids as observed in the cells of oilseeds. Moreover, organs, such as flowers and fruits, possess chromoplasts with relatively high levels of carotenoids thus give rise to the red, orange, and yellow colors. Among all types of plastids, chloroplasts are the only photosynthetically active plastids and thus are a

prerequisite for plant growth and development. In all subsequent sections the focus will be on the chloroplast.

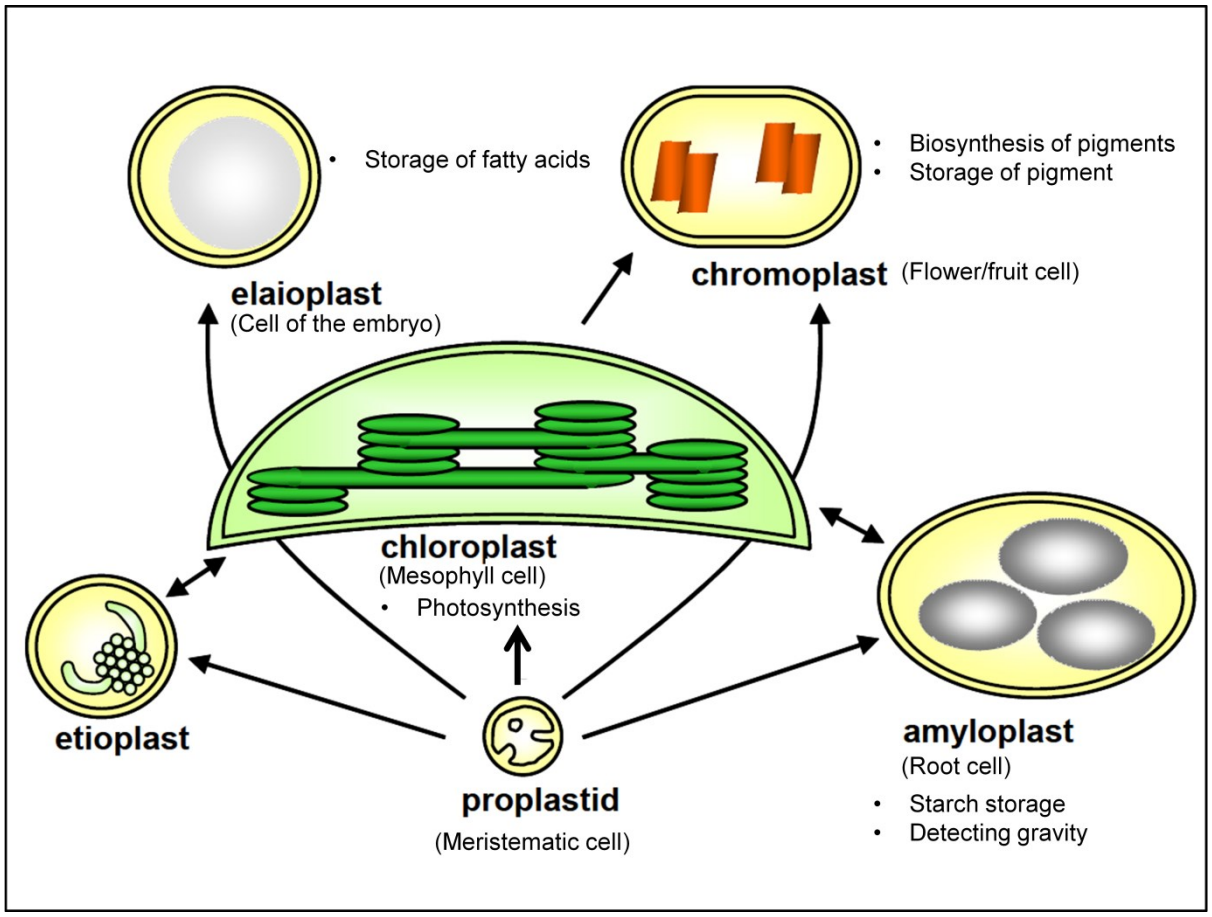

Figure 1-1: Diversity of plastid types. Plastids can differentiate into several forms, depending on their function in the cell. The functions of respective plastids were indicated as bulleted items or lists. All types of plastids derive from the undifferentiated proplastids present in meristematic cells. Etioplasts, the predecessors of chloroplasts, are formed when grown in the dark. The figure was adopted and modified from Lopez-Juez and Pyke (2005).

1.1.1 Structure and function of the chloroplast

Chloroplasts make the most prominent components of the mesophyll cells in leaf tissue of higher plants. Depending on the species, their number varies from dozens to over hundred per mesophyll cell. Each chloroplast comprises three different membranes, the double membraned envelope and the inner-most thylakoid membrane, which enclose three distinct soluble compartments (intermembrane space, stroma and thylakoid lumen). The outer and inner membrane of the envelope form the boundary to delimit the territory of a chloroplast. It builds the supporting frame for the translocon components, the TRANSLOCON AT THE OUTER

ENVELOPE MEMBRANE OF CHLOROPLASTS (TOC) and TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS (TIC) proteins, which establish pores or channels in the membrane to allow the passage of essential nucleus-encoded proteins (Li and Chiu, 2010). The internal membrane thylakoids are organized in two patterns, either forming grana which are stacks of thylakoid discs in a cylinder shape or appearing as individual lamellae in the stroma as a connector between the thylakoids. Thylakoids are sites of the light-dependent reactions of photosynthesis. The core photosynthesis related proteins or protein complexes are located at distinct locations of the thylakoids, which termed as 'lateral heterogeneity' (Jensen and Leister, 2014). PHOTOSYSTEM II (PS II) and LIGHT-HARVESTING COMPLEX II (LHC II) are limited to the grana membranes. On the contrary, PHOTOSYSTEM I (PS I) and its LIGHT-HARVESTING COMPLEX I (LHC I) as well as the ATP SYNTHASE are exclusively concentrated in the stroma lamellae. The CYTOCHROME B_6F complex is present in both types of the thylakoids. The space between thylakoids is occupied by the stroma, where carbon fixation takes place. Apart from these structural features, chloroplasts possess their own heritable information, called plastome, which is organized into complex structures, the nucleoids. On average, each nucleoid consists of 10 to 20 copies of the chloroplast genome and further RNA and various proteins (Sakai et al., 2004; Krupinska et al., 2013).

Besides photosynthesis, chloroplasts fulfill a major role in metabolism. This includes among others starch synthesis, nitrogen assimilation and fatty acid biosynthesis (Neuhaus and Emes, 2000). Moreover, it serves as a source of retrograde signaling, which is referring to the process of signaling from organelles (chloroplast and mitochondria) to modulate nuclear gene expression.

1.1.2 The chloroplast genome

Since it was demonstrated during the 1960s that chloroplasts contain their own DNA (Chun et al., 1963; Sager and Ishida, 1963), extensive studies with respect to the chloroplast genome, the plastome, established the area of 'chloroplast molecular biology'. The first physical map of chloroplast DNA was constructed for maize in 1976 (Bedbrook and Bogorad, 1976), Later on, complete chloroplast genomes were sequenced for tobacco (Shinozaki et al., 1986), liverwort (Ohyama et al., 1986), and

subsequently for rice (Hiratsuka et al., 1989) and other species (http://chloroplast.cbio.psu.edu/). To date, the sequence of the full chloroplast genome of a total of 746 eukaryotic photosynthetic organisms have been determined according to the NCBI Organelle Genome Resources Database (As of April 16, 2015. http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=2759&opt=plastid). The genome size of almost all chloroplast DNAs is within the range of 120 Kbp and 160 Kbp (Palmer, 1985; Sugiura, 1992). Normally, the chloroplast genome of higher plants, organized in a circular molecule, harbors a large single-copy (LSC) and a small single-copy region (SSC) that are separated by a large inverted repeat (IR) with a size of between 6 and 76 Kbp (Palmer, 1985; Chumley et al., 2006; Guisinger et al., 2011). Exceptions can be found in species such as pea, broad bean, alfalfa and pine, whose chloroplast DNAs do not possess the IR structure (Sugiura, 1992). It was suggested that one segment of the IR, present in the common ancestor of land plants, was lost in some legumes and conifers during evolution (Sugiura, 1992).

Chloroplast genomes, on average, contain 120 genes. These genes can be further classified into two main functional groups: one clade for the maintenance and expression of the organelle's own genes, including ribosomal RNA (rRNA) genes, transfer-RNA (tRNA) genes, ribosomal protein genes, translation factors and RNA polymerase subunits genes. The other clade contains photosynthesis associated genes, which consist of ribulose-1,5-bisphosphate carboxylase (*RuBisCO*) subunit gene(s), PS II genes, PS I genes, *cytochrome* b_6f complex genes, ATP synthase gene and *nicotinamide dehydrogenase* (*ndh*) genes (Sugiura, 1992). The majority of these genes are arranged in operons and transcribed as polycistronic precursor molecules that are subjected to splicing and nucleolytic cleavage in order to produce mature and translatable mRNAs (Stern et al., 2010; Wicke et al., 2011). For instance, the chloroplast genome of barley cv. Morex (NC_008951) contains 113 unique genes, among which, 78 are encoding proteins and 37 encode tRNAs or rRNAs (Saski et al., 2007). Taken advantage of differential RNA sequencing (dRNA-seq) (Sharma et al., 2010), transcriptome analysis of the barley chloroplast in green and white leaves of the barley mutant '*albostrians*' revealed that among the 113 chloroplast genes, eighty-nine were arranged in 20 polycistronic operons, while the remaining 24 genes were transcribed monocistronically (Zhelyazkova et al., 2012).

1.2 Variegation is an important trait for studying chloroplast biology

Variegation is a phenotype of plants developing alternating patches of green and pale green, yellow or un-colored (white) sectors in their vegetative parts (Kirk and Tilney-Bassett, 1978). It is a common phenomenon that can be observed by many ornamental plants, the most representative of such variegated examples can be found in the study of Toshoji et al. (2012). Some of the most common variegations have green and white sectors in otherwise normally green tissues and organs of the plant. Cells in the green sectors contain functional chloroplasts while cells in white sectors contain plastids that are deficient in chlorophyll and/or carotenoid pigmentation. These plastids appear to be blocked at various steps of chloroplast biogenesis because they most often lack organized internal membrane structures and/or contain only rudimentary lamellars (Rodermel, 2002).

1.2.1 Mechanisms of variegation

Variegation can arise by different mechanisms (Kirk and Tilney-Bassett, 1978) (Table 1-1). Some patterns of variegation are induced by external agents and are not heritable. For instance, chlorotic leaf sectors can be induced by partial shading, pathogen attack or nutritional deficiencies. Heritable variegations arise from mutations in nuclear or organellar (plastid and/or mitochondrial) genes that result in a failure of plastid functions to accumulate photosynthetic pigments. Depending on the compartment where the causal mutation occurs, heritable variegation can be classified into the following three categories.

Nuclear gene controlled variegation: There are two major types of nuclear gene controlled variegation depending on the genotypes of the white and green sectors. In Type I, cells of green sectors have a wild type (WT) genotype, while cells in white sectors have a mutant (MT) genotype. The three most frequent mechanisms behind Type I variegation are chimerism (Kirk and Tilney-Bassett, 1978), activity of transposable elements (TEs) (Feschotte et al., 2002) and RNA silencing (Meins et al., 2005). Chimerism refers to a state where parts of a plant meristem and thereof derived tissues exhibit a different genotype resulting in variegation, if the genotypic differences affect pigment formation or accumulation. TEs were first discovered in maize by Barbara McClintock as the genetic agents that are responsible for the sectors of altered pigmentation on mutant kernels (McClintock, 1951). In brief, TE

activity generates variegated plants when insertion of a transposon interrupts a nuclear gene required for normal chloroplast biogenesis (white sectors), while element excision may reconstitute WT gene expression (green sectors). Some of the well-known examples of mutable alleles of genes crucial for chloroplast biogenesis include maize *bundle sheath defective 1* (*bsd1*) (Hall et al., 1998), *defective chloroplasts and leaves* (*dcl*) in tomato (Keddie et al., 1996) and *pale-yellow-leaf variegated* (*pyl-v*) in rice (Tsugane et al., 2006). RNA silencing refers to a phenomenon of variable gene expression controlled by small RNAs derived from double-stranded RNAs (dsRNAs) or stem-loop RNA precursors (Meins et al., 2005). These RNAs guide the cleavage of target gene RNAs, block their expression or induce methylation of target genes. As a consequence, it can produce variegation when a nuclear gene required for chloroplast biogenesis is silenced in some cells but not in others. A well-known example of this is the production of variegated flowers in transgenic petunia that contained an antisense *chalcone synthase* gene (van der Krol et al., 1988); and *IspH* transgene-induced gene silencing is a recent example of leaf variegation in Arabidopsis (Hsieh and Goodman, 2005).

In Type II variegation cells (regardless of green or white) have a uniform MT genotype, but the mutant phenotype is expressed only in a subset of cells (white sectors). Variegation of this sort is typically induced by nuclear recessive genes. One of the earliest examples is *iojap*, a recessive striped mutant of maize affected in the development of plastids (Walbot and Coe, 1979). The plastids in respective white sectors did not contain 70S ribosomes and were permanently defective because the translation machinery in plastids was totally blocked. Hence, the *iojap*-affected plastids were irreversibly defective even if the nuclear gene function was restored. Cloning of the *iojap* gene revealed it to encode a component of the 50S subunits of the plastid ribosome (Han et al., 1992); however, the exact molecular mechanism leading to variegation in this mutant remains unclear. Another well-known Type II variegation mutant is *albostrians* of barley, which also controlled by nuclear recessive gene (Hagemann and Scholz, 1962; Hess et al., 1994a).

Plastid gene controlled variegation: In addition to variegation controlled by nuclear genes, variegation can also derive from mutations of the plastid genome, i.e. plastome mutations. Plastome mutations can arise spontaneously, by artificial mutagenesis, by action of a nucleus-encoded chloroplast mutator or by chloroplast

transformation (Börner and Sears, 1986; Maliga, 2004). Most of the plastome mutations of higher plants originated spontaneously. Mutation rates can be increased by ionizing irradiation or by treatment with mutagenic chemicals. In addition to plastome mutations caused by the above mentioned external factors, chloroplast mutators, i.e. nuclear genes can also cause mutations in the chloroplast genome. Because chloroplast genomes consist of multiple copies, it is thought that the chloroplast mutator lines, as well as other plastome mutants, are variegated because mutant and normal plastid chromosomes, following replication, sort out to form clones of plastids and subsequent cells contain either all-normal plastid DNAs (green sectors) or all-mutant plastid DNAs (white sectors). Cells containing a single type of plastid DNA are termed 'homoplasmic', whereas ones with different types are termed 'heteroplasmic'. In addition, plastome mutations can be caused by chloroplast transformation. The principle of chloroplast transformation is based on integrating an antibiotic-resistance expression cassette into the plastome via homologous recombination (Svab et al., 1990). The plastomes primarily code for core components of the photosynthetic apparatus and for proteins involved in plastid gene expression, it might be anticipated that mutations in many plastid genes would give rise to defective plastids. Similar to plastome mutators, variegations arise when these genomes sort out to form homoplasmic clones of plastids and cells.

Mitochondrial gene controlled variegation: In addition to variegations that are caused by the induction of permanently-defective plastids by nuclear plastome-mutator genes, variegations can also be caused by mutations in nuclear genes that generate permanently-defective mitochondria. Notable examples include the *nonchromosomal stripe* (*NCS*) mutants of maize (Newton and Coe, 1986) and *chloroplast mutator* of Arabidopsis (Sakamoto et al., 1996). These mutants are variegated because the abnormal mitochondria secondarily affect the function of the plastids in the cell. Hence, these sorts of variegation are Mendelian-inherited, but the defective mitochondria are inherited maternally (or bi-parentally).

1.2.2 Genetic factors underlying variegation

The variegation mutants provide an excellent system to study mechanisms of chloroplast biogenesis. Variegated mutants of Arabidopsis were considered an optimal tool to unveil mechanisms of chloroplast biogenesis, since Arabidopsis is a

model plant species with the most comprehensive molecular tools, resources and knowledge already in place. To date, several genes controlling variegated phenotypes in Arabidopsis were cloned, such as *chloroplast mutator* (*chm*) (Abdelnoor et al., 2003), *immutant* (*im*) (Carol et al., 1999; Wu et al., 1999), *yellow variegated 1* (*var1*) (Sakamoto et al., 2002), *yellow variegated 2* (*var2*) (Chen et al., 2000) and *yellow variegated 3* (*var3*) (Naested et al., 2004). Besides these genes that are directly linked to the variegation phenotype, a handful of *var2* genetic suppressor genes that could compensate the variegated phenotype were identified, such as the chloroplastic Hsp100 chaperone gene *clpC2* (Park and Rodermel, 2004), *fu-gaeri1* (*fug1*) and *snowy cotyledon 1* (*soc1*) (Miura et al., 2007), *suppression of variegation 1* (*svr1*) and *suppression of variegation 2* (*svr2*) (Yu et al., 2008), G protein α subunit gene *GAP1* (Zhang et al., 2009b), and *suppression of variegation 7* (*svr7*) (Liu et al., 2010) providing additional insights into mechanisms underlying variegation.

Besides the mutants mentioned above for Arabidopsis, the most prominent and well-studied variegation mutants are the above mentioned *iojap* mutant of maize and the *albostrians* mutant of barley. Although the *albostrians* mutant is well characterized biologically and biochemically, the causal gene has not been identified and cloned so far.

Table 1-1: Summary of mechanisms underlying variegation.

Mechanism	Description	Inheritability	Genotype (White *vs* Green)	Determinant	Examples
Chimerism	Tissues derive from different histological regions of a plant meristem.	Yes	Distinct	Nuclear gene	-
Transposable element activity	Insertion of transposon interrupts a nuclear gene required for normal chloroplast biogenesis.	Yes	Distinct	Nuclear gene	Maize *bsd1*; Tomato *dcl1*; Rice *pyl-v*
RNA silencing	The RNAs guide the cleavage of target gene RNAs, block their translation or induce methylation of target genes.	Yes	Distinct	Nuclear gene	Arabidopsis *IspH*
Nuclear genome mutations	Nuclear gene defect blocks chloroplast biogenesis.	Yes	Uniform	Nuclear gene	Arabidopsis v*ar2*; Maize *iojap*; Barley *as*
Plastome mutators	Nuclear genes that cause mutations in chloroplast DNA.	Yes	Distinct	Plastid gene	Evening primose *pm*
Plastome mutations	Arise spontaneously or by chemical treatment or by chloroplast transformation.	Yes	Distinct	Plastid gene	-
Mitochondrial genome mutations	Mitochondrial defect leads to lesions in chloroplast biogenesis.	Yes	Distinct	Mitochondrial gene	*NCS* mutants of Maize/Tobacoo/Tomato
Plastid-nucleus incompatibility	Developmental disturbances accompanied by pigment deficiencies.	Yes	Distinct	Nuclear gene & plastid gene	Evening primrose
External regents	Variegations induced by preferential shading, pathogen attack and nutritional deficiencies.	No	-	-	-

1.3 *albostrians* is a classical variegation mutant of barley

1.3.1 Background of the *albostrians* mutant

The *albostrians* mutant originated from the two-rowed spring barley variety ‘Haisa’ by means of X-ray irradiation (Hagemann and Scholz, 1962). *Albino*, green-white striped and pure green seedlings can be observed when growing progeny of a fully green but homozygous *albostrians* mutant (*as*/*as*) (Figure 1-2). This pattern of segregation is following a ratio of 1:8:1 (green / striped / *albino*). Based on the genetic segregation analysis in F_2 generation obtained by crossing the mutant to a wild-type genotype and self pollinating the obtained F_1, however, it could be shown that the phenotype was caused by a single recessive, nucleus-encoded gene (Hagemann and Scholz, 1962). In contrast, the inheritance of the chlorophyll deficiency is following a purely maternal pattern, as was revealed by reciprocal crosses between Haisa and the mutant line M4205 (Figure 1-3) (Hagemann and Scholz, 1962). The *albostrians* mutant was therefore initially considered as a nuclear gene induced plastome mutation, i.e. a mutation of chloroplast DNA induced by action of a nuclear mutator gene (Hagemann and Scholz, 1962).

Figure 1-2: Examples of variegation of coloration in homozygous *albostrians* mutant plants.
The leaves exhibited here were collected from different seedlings of the mutant line M4205.

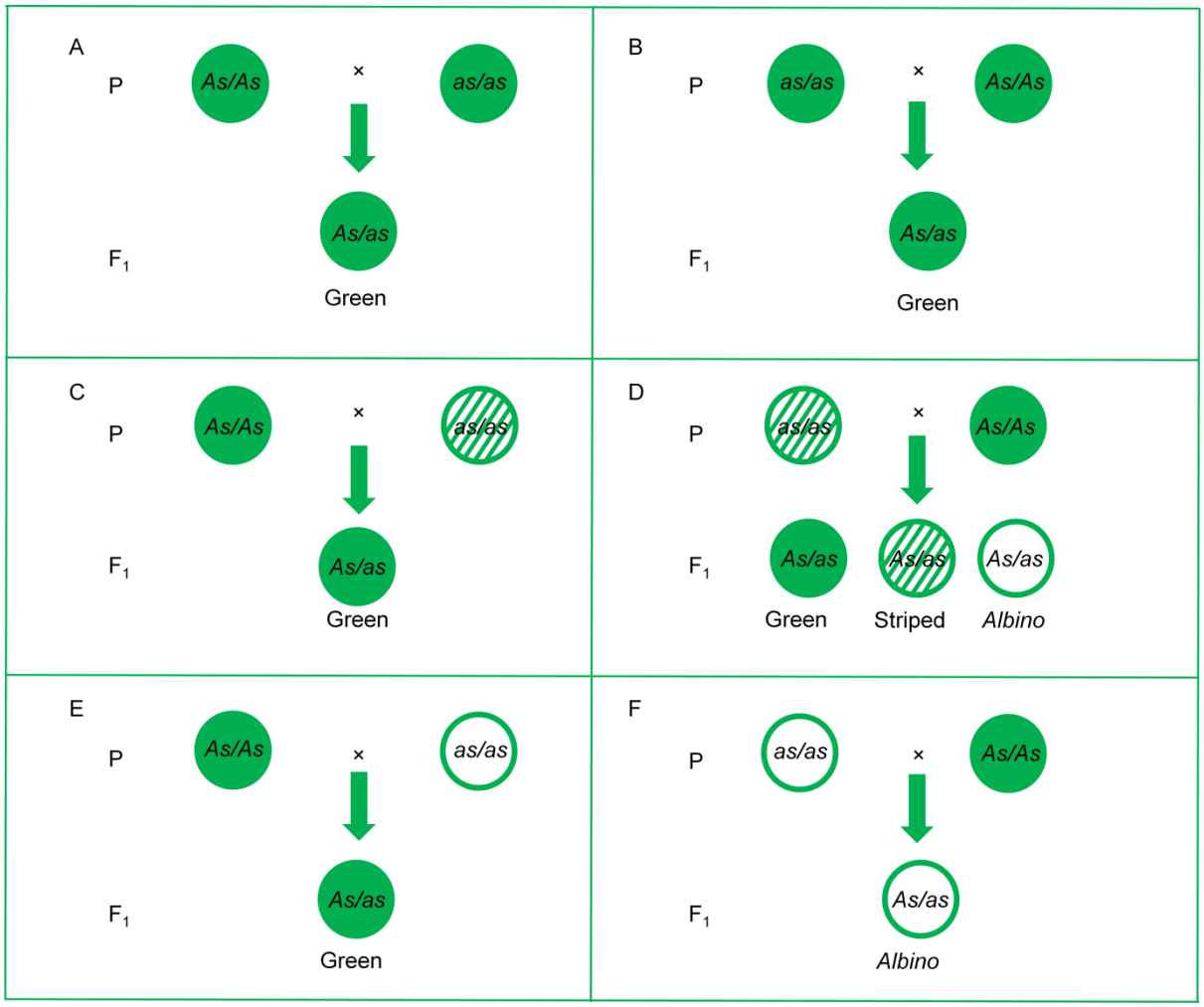

Figure 1-3: Pattern of inheritance of the *albostrians* mutant phenotype based on chloroplast constitution. The inheritance of leaf color variegation is depending on the chloroplast population in the maternal plant as can be shown in offspring of reciprocal crosses. If WT (*As*/*As*) or a green spike of M4205 (*as/as*) is used as the female parent, all F_1 plants are green (A, B, C, E). A striped spike as female parent given green, striped and *albino* F_1 plants (D), and a white spike as female parent produces only *albino* F_1 plants because the plastid aberration is irreversible and cannot be rescued (F). The figure is drawn on the basis of the crossing experiments done by Hagemann and Scholz (1962).

1.3.2 Status of research towards characterizing the *albostrians* mutant

The most prominent characteristic of the *albostrians* mutant is the absence of 70S ribosomes in plastids of white sectors. This was supported by five lines of evidence: 1) neither 16S nor 23S rRNA could be found in preparations of total RNA from white leaves; 2) no ribosomes were observed by transmission electron microscopy in the white plastids; 3) the CF1-ATPase consisting of nuclear and plastid encoded subunits was lacking in the white *albostrians* leaves; 4) ribosomal protein L2 was never

detected by means of western blot in extracts from white *albostrians* seedlings; 5) white leaves lack RUBISCO the large subunit of which is synthesized on chloroplast ribosomes (Börner et al., 1976; Reichenbächer et al., 1978; Börner et al., 1979; Hess et al., 1993; Hess et al., 1994a). Since the functional ribosome is an essential apparatus involved in the process of translation, theoretically, translation of all chloroplast genes coding for proteins is entirely abolished in the white sectors of the mutant. Because of this aspect, the ribosome-free plastids of the *albostrians* mutant were considered an excellent system for studying regulatory interactions between the three DNA-containing compartments, for instance, the impact of protein synthesis deficiency in plastids on gene expression in nucleus and mitochondria as well as on plastids themselves (Hess et al., 1994a).

The first significant achievement reached by studying the *albostrians* mutant was the first evidence for the existence of plastid to nucleus signals (Bradbeer and Börner, 1978; Bradbeer et al., 1979), now known as retrograde signaling (Pesaresi et al., 2007). Initially, instead of using the term ‘retrograde signals’, a ‘chloroplast control principle’ (Hagemann and Börner, 1978) was postulated. Plastids were thought to influence nucleo-cytoplasmic gene expression. This could be the result of affecting the accumulation of nuclear gene-encoded chloroplast polypeptides, such as LIGHT-HARVESTING CHLOROPHYLL A/B-BINDING PROTEIN (LHCP), the small subunit of RUBISCO (Hagemann and Börner, 1978), the GLYCERALDEHYDE PHOSPHATE DEHYDROGENASE (NADP+) and PHOSPHORIBULOKINASE (Bradbeer and Börner, 1978). All these proteins are known to be of nucleocytoplasmic origin and were drastically reduced to the limit of detection in their activities and/or quantities in white leaves of the *albostrians* mutant. Plastids were also shown to affect the accumulation of mRNAs transcribed from nuclear genes, or the accumulation of the non-chloroplast enzyme NITRATE REDUCTASE (Börner, 1986; Börner et al., 1986). The proposed ‘chloroplast control principle’ was in agreement with the ‘multi-subunit completion principle’ (Ellis, 1977) supporting the idea that proteins within the chloroplast are synthesized on both plastid and cytoplasmic ribosomes. However, the ‘chloroplast control principle’ was not compatible with the ‘cytoplasmic control principle’ postulated by Ellis who insisted on the requirement that cytoplasmic products control organellar protein synthesis, but not the other way around (Ellis, 1977). On the contrary, the ‘chloroplast control principle’ proposed that the

nucleocytoplasmic compartment and the plastid behave as cooperation partners, i.e. plastid gene expression is under the control of nuclear genes and vice versa (Hagemann and Börner, 1978). Interestingly, the 'signaling factor' was initially postulated to be RNA transcribed from the plastid genome (Bradbeer and Börner, 1978); a new hypothesis was subsequently assumed by accounting the intermediates of chlorophyll biogenesis serving as the connection between nuclear gene expression and chloroplast development (Hess et al., 1992; Hess et al., 1994b). Furthermore, it was considered the initial steps of Mg-porphyrin biosynthesis to contribute to plastid-derived signaling towards the nucleus (Yaronskaya et al., 2003). The latter hypothetical scenario was supported by the finding that Mg-PROTOPORPHYRIN IX acts as a negative regulator of photosynthesis gene expression in the nucleus and the chloroplast (Strand et al., 2003; Ankele et al., 2007). Besides the plastid to nucleus retrograde signaling, it is noteworthy to mention that studies on *albostrians* mutant of barley provided also for the first time evidence for an influence of the plastids/chloroplasts on the expression of mitochondrial genes and mitochondrial gene copy numbers (Hedtke et al., 1999).

The second milestone reached by investigating the *albostrians* mutant was finding the first evidence for the existence of two different chloroplast RNA polymerase systems, i.e. plastid-encoded plastid RNA polymerase (PEP) and nucleus-encoded plastid RNA polymerase (NEP) (Hess et al., 1993). As early as in 1970, the idea was proposed for maize where chloroplasts contained two different types of DNA-dependent RNA polymerase (Bogorad and Woodcock, 1970). Studies with respect to the *albostrians* mutant clearly revealed that the plastid genes, *RNA polymerase B, C1 and C2* (*rpoB/C1/C2*) as well as 40S *ribosome protein S15* (*rps15*), were transcribed despite the lack of PEP subunits in the ribosome-deficient plastids. The *rps15* gene and *rpo* genes showed high gene expression in ribosome-deficient plastids, in contrast to photosynthetic genes. On the contrary, the functional chloroplast contained abundant transcripts of the photosynthetic genes but not of the *rpo* genes. It was speculated that PEP has the preference for expression of the photosynthetic and bioenergetic genes, while, the NEP has preference for transcription of the housekeeping genes (Hess et al., 1993). This assumption was consistent with observations in tobacco and Arabidopsis showing that genes of the two photosystems completely relied on PEP transcription, while, transcription of

housekeeping genes was depending on both polymerase systems; with the exception, that the *rpoB* operon was transcribed exclusively from NEP promoters (Hajdukiewicz et al., 1997; Swiatecka-Hagenbruch et al., 2007). In order to reach a comprehensive understanding of the division of labor of the two polymerases, the chloroplast primary transcriptome of the *albino* and green *albostrians* leaves were investigated revealing that PEP plays the dominant role in transcription of the genes in mature chloroplasts (Zhelyazkova et al., 2012). In addition, it provided evidence that NEP is also involved in transcription of genes for the photosynthetic apparatus, such as *rbcL*, *psbB*, *psbD* and *psbM*, which were previously considered to be transcribed only by PEP (Hajdukiewicz et al., 1997; Swiatecka-Hagenbruch et al., 2007)

Overall, the *albostrians* mutant of barley served as a versatile genetic material to address various biological questions mainly related to chloroplast biogenesis. Nevertheless, the causal gene underlying the variegated phenotype still remained unknown. From this perspective, identification of the gene would provide a genetic basis towards a better understanding of the mechanisms underlying *albostrians* based variegation and maybe variegation in general.

1.4 Current genomic resources of barley

In addition to its importance for agriculture, barley has been a genetic model organism since early of last century. However, the major impediment to make a full exploitation of the available genetic resources of barley has been absence of genomic resources. In this regard, the global barley community increased their efforts on establishing a variety of versatile genomic resources to underpin a better utilization of the existing genetic resources in basic and applied research.

The genomic resources of barley were improved during the past decade, first of all, by the development of several high density genetic maps (linkage map & consensus map) from several individual bi-parental or doubled haploid (DH) mapping populations (Rostoks et al., 2005; Wenzl et al., 2006; Marcel et al., 2007; Stein et al., 2007; Varshney et al., 2007; Potokina et al., 2008; Close et al., 2009; Sato et al., 2009a). In order to further increase marker density to promote genome-wide association study (GWAS), a 9K iSelect chip, containing 7,864 selected single nucleotide polymorphisms (SNPs), was designed based on a total of 240,119 SNPs

data-mined from transcriptome sequencing data obtained of pooled RNA samples of 10 different barley cultivars (Comadran et al., 2012). In addition, combined with next generation sequencing (NGS) technology, the newly developed genotyping-by-sequencing (GBS) approach, which integrated SNPs discovery and genotyping in one step, made marker development much more efficient. Compared to the previously published genetic maps, the amount of markers produced via GBS approach could be increased by a factor of 10 (Poland et al., 2012; Poland and Rife, 2012). This has greatly facilitated marker development during gene identification by map-based cloning. Regarding to transcriptome analysis, a 22K Barley1 GeneChip based on an expressed sequence tags (EST) database (http://harvest-web.org/hweb/pickassy.wc) containing 350,000 sequences from 84 different RNA sources was established (Close et al., 2004). Subsequently, quite a few barley sequence resources were generated, such as the 22,651 full-length cDNA (FLcDNA) sequences derived from barley cultivar 'Haruna Nijo' (Sato et al., 2009b; Matsumoto et al., 2011). In addition, five bacterial artificial chromosome (BAC) libraries of barley cultivar 'Morex', representing 25-fold haploid genome coverage, were generated for map-based cloning and physical map construction (Schulte et al., 2011). Before emerging of the physical map of barley, as a proof of principle, a virtual linear gene order of the barley chromosomal 1H was established by an integrated cytogenetics, molecular genetics, and bioinformatics approach and subsequently successfully applied for the additional six chromosomes (Mayer et al., 2009; Mayer et al., 2011). Recently, a milestone achieved by the International Barley Genome Sequencing Consortium (IBSC) was the released physical, genetic and functional sequence assembly of the barley genome, which serves now as an invaluable resource for genome-assisted gene isolation (International Barley Genome Sequencing Consortium, 2012). Since then, an improved version of the physical map of barley was achieved by using a newly developed population sequencing (POPSEQ) anchoring approach with more than two million SNPs as the genetic framework (Mascher et al., 2013a; Ariyadasa et al., 2014). In addition, the recently established exome capture platform, with 61.6 Mb enriched coding sequence, enables for reduced complexity re-sequencing of the barley genome, and has been proven as a powerful tool for gene isolation and diversity analysis (Mascher et al., 2013b; Mascher et al., 2014).

Barley genomic resources can be accessed via a set of web-based tools and platforms such as IPK Blast Server (International Barley Genome Sequencing Consortium, 2012), MIPS PlantsDB (Spannagl et al., 2007; Nussbaumer et al., 2013), BarleyBase (Shen et al., 2005; Wise et al., 2007), GrainGenes (Matthews et al., 2003; Carollo et al., 2005; O'Sullivan, 2007), HarvEST (Close et al., 2004), Gramene (Youens-Clark et al., 2011; Monaco et al., 2014), and Ensembl Plants (Bolser et al., 2015). The accumulated genomic resources of barley provide a foundation for cloning of biologically and/or agronomically important genes.

1.5 Map-based cloning in barley

In general, the identification of a particular gene as a basis for analysis of its function can be addressed by forward and reverse genetics approaches. Forward genetics refers to the strategy starting from a given phenotype working towards the identification of the underlying functional gene. In contrast, for an attempt of reverse genetics a gene has been identified and methods like mutagenesis or transgenesis are used to discover a phenotype caused by the defective gene. Typically, mutant phenotypes were known long before the underlying genes could be identified. In barley, often only genetic information is available for genes underlying agronomical traits or for natural or induced mutants. Here, forward genetics by map-based cloning is the most promising approach towards molecular isolation of the underlying functional genes (Krattinger et al., 2009). Map-based cloning refers to a strategy that aims at the genetic and physical delimitation of the region in the genome of a given species that must contain the functional gene. Basically, map-based cloning contains the following three steps: 1) allocate the target gene to a defined interval by genetic mapping, 2) physical mapping, identification and isolation of the candidate gene, and 3) confirmation of the identified candidate gene by functional analysis, such as functional complementation or mutant analysis (Wing et al., 1994; Stein and Graner, 2005). The newly generated genomic resources of barley have made it possible to exploit extensively the available genetic resources. As of to date, a series of barley genes with important biological functions was isolated through map-based cloning; among others genes conferring resistance to fungal and viral pathogens (*mlo*, Buschges et al., 1997; Simons et al., 1997; *Mla6*, Halterman et al., 2001; *Rpg1*, Brueggeman et al., 2002; *rym4/rym5*, Stein et al., 2005; *rym11*, Yang et al., 2014), genes affecting plant architecture (*Vrs1*, Komatsuda et al., 2007; *Nud*, Taketa et al.,

2008; *Vrs4*, Koppolu et al., 2013), and genes controlling flowering time (*Ppd-H1*, Turner et al., 2005; *Mat-a*, Zakhrabekova et al., 2012).

1.6 The aims of the study

Previous studies of the *albostrians* mutant of barley mainly focused on elucidating physiological and regulatory interactions between plastids and the other DNA containing compartments in plant cells. This included also questions regarding the role of the plastid and nuclear genomes in the control of plastid metabolism. These studies remained limited due to the fact that the functional gene underlying the *albostrians* phenotype remained unknown. The main objective of the present study was to clone the barley gene *albostrians* through map-based cloning supported by recent advances in the generation of barley genomic resources. The essential work to reach this goal included the following aspects:

- Genetic mapping of the *albostrians* gene.
- Physical mapping of the gene *albostrians* by taking advantage of the emerging physical map of the barley genome.
- Identification of a candidate gene by sequence analysis of the physically delimited target region in the barley genome.
- Functional analysis of the identified candidate gene by mutant and transgenic analysis.
- Analyzing the subcellular localization of the protein encoded by the *albostrians* candidate gene.

Overall, the isolation of the gene *albostrians* would provide a critical achievement for reaching a better understanding of early steps in chloroplast maturation.

2 Material and Methods

2.1 Plant material

Two F_2 mapping populations, MM4205 and BM4205, were constructed by crossing the spring barley cultivars (cv.) 'Morex' and 'Barke' to the mutant line M4205 (Hagemann and Scholz, 1962), respectively. Initially, 91 genotypes from each population were selected for low-resolution mapping the *albostrians* gene. Subsequently, only the MM4205 population was further maintained for fine mapping the *albostrians* gene due to its higher number of polymorphic markers compared to the BM4205 mapping population. The mapping population was first increased to 1051 F_2 genotypes by growing additional 960 individuals. In a second step of increasing the mapping population size, 384 homozygous *as/as* mutant plants (exhibiting *albino* or variegated phenotype) were selected phenotypically from a total of 1920 F_2 plants. These plants were further analyzed with flanking markers for the identification of recombinants at the target locus. This strategy allowed to save time since it was not necessary to wait for phenotyping in F_3 families in order to determine the genotype at the target locus.

Seeds were germinated in 96-well plates containing soil. Then, the selected recombinants were transferred into individual pots (14 cm diameter) and grown under greenhouse conditions with a temperature cycle of 20°C / 15°C (day/night) and a photoperiod of 16h light / 8h dark. The exact number of plants grown from each F_2 family for the genetic mapping analysis is summarized in Appendix Table 1 and a schematic illustration of population development is given in Figure 2-1.

For mutant analysis, a TILLING population (Gottwald et al., 2009) derived from barley cultivar 'Barke' was screened for the genomic region of the *albostrians* gene. The plant families carrying lesions were selected for analysis. In addition, a set of 28 *albino* plant families were obtained from the Nordic Genetic Resource Center (NGRC, NordGen) (Table 2-1) with the aim to identify independent mutant alleles of the gene *albostrians*. All plants were nurtured in the greenhouse as mentioned above.

Table 2-1: List of *albino* plant families ordered from NGRC.

Mutant Family ID	Mother Cultivar	Mutagen	Year
alb-a.60	Bonus	diepoxybutane	1957
alb-b.76	Bonus	1,4-bis(methane-sulfonyloxy) butane	1956
alb-c.7	Gull	spontaneous	Ca. 1925
alb-d.13	Bonus	neutrons	1954
alb-e.16	Bonus	x-rays	1954
alb-f.17	Bonus	x-rays	1954
alb-g.18	Bonus	x-rays	1954
alb-h.19	Bonus	x-rays	1954
alb-i.22	Maja	neutrons	1955
alb-j.24	Maja	x-rays	1955
alb-k.25	Maja	neutrons	1955
alb-m.27	Bonus	x-rays	1955
alb-n.28	Bonus	x-rays	1955
alb-o.32	Bonus	Triethylene-melamine	1956
alb-p.41	Bonus	Alpha-rays	1951
alb-q.45	Bonus	x-rays	1952
alb-r.52	Bonus	neutrons	1952
alb-s.59	Bonus	ethyleneimine	1957
alb-t.66	Bonus	Triethylene-melamine	1956
alb-u.78	Bonus	p-N-di (b-chloroethyl) phenylalanine	1956
alb-v.81	Bonus	Ethylene oxide	1956
alb-x.89	Bonus	Ethylene oxide	1956
alb-y.90	Bonus	Ethylene oxide	1956
alb-z.95	Bonus	nebularine	1957
alb-zb.113	Bonus	ethyleneimine	1958
alb-zc.122	Bonus	ethyleneimine	1958
alb-zd.133	Bonus	ethyleneimine	1957
alb-ze.134	Bonus	ethyleneimine	1957

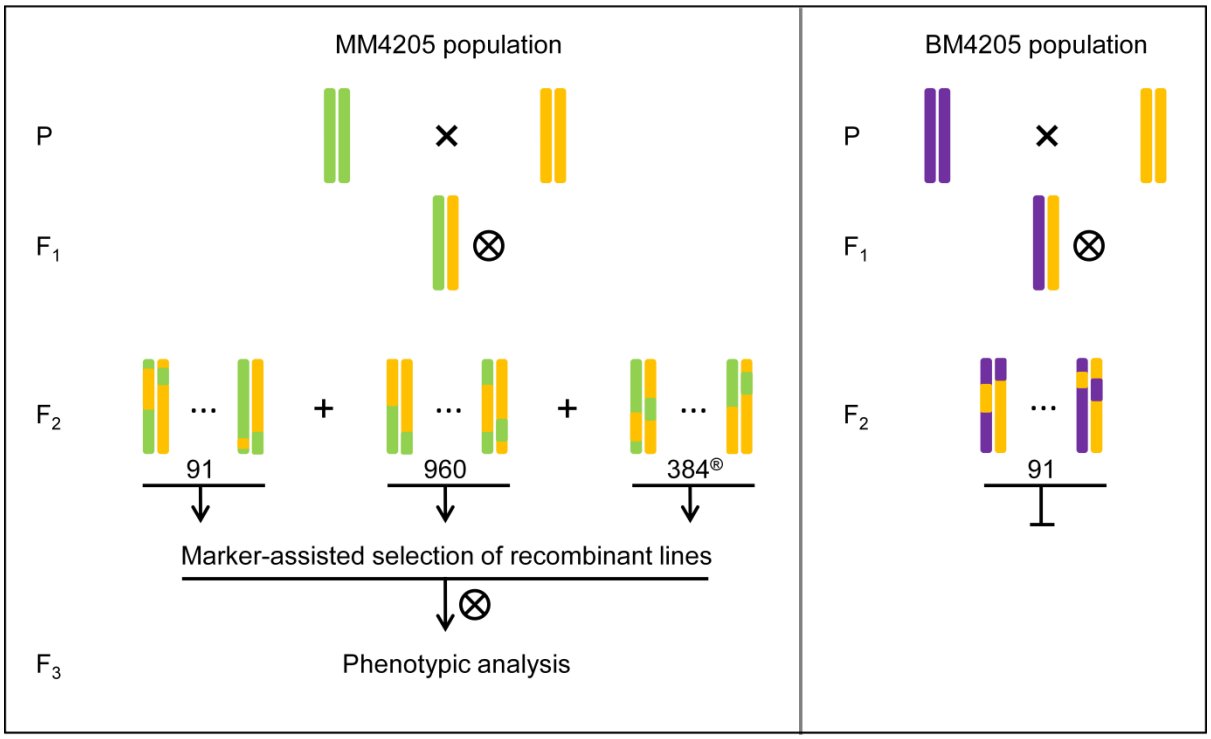

Figure 2-1: Schematic overview of population development and genetic mapping. Two F_2 populations MM4205 and BM4205 were constructed for mapping the gene *albostrians*. Low-resolution mapping was performed with 91 F_2 genotypes from each of the population. High-resolution mapping was initiated after no further recombination could be detected between the flanking markers within the 91 genotypes. Only MM4205 mapping population was further used for high-resolution mapping due to higher level of polymorphism compared to the BM4205 mapping population. The population was first increased with additional 960 genotypes and subsequently enlarged by analyzing an additional set of 384 variegated or *albino* seedlings. The latter population was selected from 1920 F_2 seedlings. The finally selected recombinants were analyzed phenotypically in F_3 to confirm genotype of the *albostrians* gene.

2.2 Phenotyping

Phenotyping of the F_2 mapping population was performed at 1-, 2- and 3-leaf seedling stages. The phenotype of the seedlings was classified into three categories: green, variegated and *albino*. Green phenotype defined all the three leaves of each seedling were purely green; variegated phenotype defined green-white striped pattern can be observed on all or any of the three leaves; and *albino* phenotype defined the seedling was completely white. The variegated and *albino* plants are homozygous for the mutant allele (*as/as*). Green seedlings, however, can either be wild-type (*As/As*), heterozygous (*As/as*) or homozygous mutant (*as/as*) at the *albostrians* gene locus. Therefore, the *albostrians* genotype was determined by phenotyping 30 seedlings of each F_3 family. A wild-type F_2 would produce 100%

green progeny, progeny of the heterozygous F_2 would follow a Mendelian segregation [Green : (variegated & *albino*) = 3:1], while the F_2 homozygous green mutant would segregated into 10% green, 80% variegated, and 10% *albino* seedlings in F_3 generation. Genotype of the *albostrians* allele of each F_2 was scored on the basis of the F_3 segregation (Appendix Table 2).

2.3 DNA isolation

Total DNA was isolated from the second leaf of 14-day-old seedlings using a published protocol (Doyle and Doyle, 1990). Plant material, frozen in liquid nitrogen, was homogenized using Mixer Mill MM400 (Retsch GmbH, Haan, Germany). Four hundred and fifty microliter extraction buffer [20 mM Ethylenediaminetetraacetic acid (EDTA), pH=8.0; 1.4 M NaCl; 100 mM Tris-HCl, pH=8.0; 53 mM $Na_2S_2O_5$; 55 mM cetyltrimethylammonium bromide (CTAB); 0.28 M ß-Mercaptoethanol; warmed to 65°C] were added to 50-100 mg of homogenized leaf material and incubated 30 min at 65°C in a water bath. Subsequently, 430 µl CIA (Chloroform:Isoamyl alcohol = 24:1) were added to each sample, followed by mixing on an overhead shaker (Keison Products, Chelmsford, England). The supernatant obtained after centrifugation (2,240 x *g* for 10 min at 4°C) was transferred to new 8-well strip tubes (Qiagen Hilden, Germany). This step was repeated once after adding 380 µl CIA to each sample. Aqueous supernatants were transferred to new 8-well strip tubes containing 45 µl acetate-mix (3M sodium acetate:10M ammonium acetate = 6:5 v/v). Additionally, 260 µl pre-chilled isopropanol were added to each sample, followed by mixing to precipitate the DNA. Samples were centrifuged (2,240 x *g* for 30 min at 4°C) and the obtained pellets were washed with 500 µl of 70% (v/v) ethanol. To dissolve the DNA, 100 µl TE-buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA, pH 8.0) and 2 µl RNase (0.1 mg/ml; Carl Roth GmbH, Karlsruhe, Germany) were added to each sample and incubated at room temperature until DNA was completely dissolved. Chloroform and ß-Mercaptoethanol are toxic through inhalation of vapors thus all manipulations involving chloroform and ß-Mercaptoethanol had to be performed in a chemical fume hood. The quantity and quality of the obtained DNA was determined by help of a NanoDrop Spectrophotometer (Thermo Scientific, Wilmington, USA) following manufacturer's instructions. Concentration was determined based on sample absorbance at 260 nm, and purity was checked by the ratio of sample absorbance at 260 and 280 nm ($A_{260}/A_{280} \geq 1.8$ indicates 'pure'

nucleic acids). The DNA was kept at -20°C for long term storage, after taking an aliquot for preparing a diluted sample [40-80 ng/µl] for immediate use in downstream PCR applications.

2.4 Polymerase chain reaction (PCR)

DNA amplification reactions were performed in a total volume of 20 µl containing 40 ng of template DNA, 4 mM of dNTPs, 1 µl each of 5 µM forward and reverse primer, 0.5 units of HotStarTaq DNA polymerase (Qiagen, Düsseldorf, Germany) and 2 µl of 10x PCR buffer (100 mM Tris-HCl, pH 8.3; 500 mM KCl; 15 mM $MgCl_2$; 0.01% gelatin). Touch-down PCR program was used with a GeneAmp 9700 thermal cycler (Life Technologies GmbH, Darmstadt, Germany): first denaturation at 94°C for 15 min followed by 5 cycles at 94°C for 30 s, annealing at 65°C to 60°C (-1°C/cycle) for 30 s, extension 1 min at 72°C, and then proceeded with 40 cycles 94°C for 30 s, 60°C for 30 s, 72°C 1 min, and followed by a final extension at 72°C for 10 min.

2.5 Agarose gel electrophoresis analysis

The PCR-amplified products were resolved by gel-electrophoresis. 1% (w/v) agarose gel was prepared by melting 1 g of UltraPure™ Agarose (Invitrogen GmbH, Darmstadt, Germany) in 100 ml of 1x TBE buffer (89 mM Tris-borate, pH 8.3; 2 mM Na_2EDTA) (SAMBROOK et al., 1989). Along with the PCR products, DNA size standard 1 kb DNA Ladder (Fermentas GmbH, St. Leon-Rot, Germany) was loaded in parallel for fragment size estimation. The gel was running in an electrophoresis chamber (Bio-Rad Laboratories GmbH, München, Germany) immersed with 1x TBE buffer. DNA molecules were visualized by staining gels with ethidium bromide [0.5 µg/ml] and photographed under UV light using the BioDocAnalyze Gel-analyze System (Biometra GmbH, Göttingen, Germany)

2.6 RNA isolation

All the procedures for isolating and handling of RNA were performed under RNase-free conditions – a specified RNA isolation workbench was cleaned with RNase AWAY™ surface decontaminant (Thermo Scientific, Wilmington, USA), and all the solutions were prepared with 0.1% (v/v) Diethylpyrocarbonate-treated water (DEPC-H_2O, and therefore RNase-free). RNA isolation was performed using a TRIzol® reagent (Invitrogen GmbH, Darmstadt, Germany) following manufacturer's

instructions. In brief, 1 ml TRIzol® reagent was added to 50-100 mg of frozen (liquid nitrogen) leaf material. The sample was immediately homogenized by using a Vortex Mixer (Keison Products, Chelmsford, England) and incubated at room temperature for 5 min. Then 200 µl chloroform ($CHCl_3$) was added and the sample mixed by inverting the tube multiple times and kept at room temperature for another 2-3 min. The supernatant obtained after centrifugation (10,000 x *g* for 20 min at 4°C) was transferred to a new RNase-free tube, mixed with 500 µl isopropanol, followed by inverting the tubes 4-6 times and incubating at room temperature for 10 min. Samples were centrifuged again (13,800 x *g* for 10 min at 4°C) and the derived RNA pellet was washed with 1 ml of 75% ethanol. The air dried RNA pellet was re-suspended in 0.1% (v/v) DEPC-H_2O and stored at -70°C for downstream applications. The concentration of the obtained RNA was determined by help of a Qubit® 2.0 Fluorometer (Life Technologies GmbH, Darmstadt, Germany) according to manufacturer's instructions. In contrast to the UV-absorbance method, the Qubit fluorometer used RNA standards to determine the relationship between RNA concentration and fluorescence.

2.7 Formaldehyde agarose gel electrophoresis

Formaldehyde agarose gel electrophoresis was used to check the integrity of the extracted RNA. Electrophoresis was performed under RNase-free conditions - the electrophoresis chamber and comb were washed with 0.1% (v/v) DEPC- H_2O and the agarose gel, which contained 2% (w/v) agarose, 1X 3-(N-morpholino) propanesulfonic acid (MOPS) buffer and 6.29% (v/v) formaldehyde, was prepared with 0.1% (v/v) DEPC- H_2O. The RNA sample (1-5 µg) was mixed with formaldehyde loading dye, which contained 25 µl formamide (Carl Roth GmbH, Karlsruhe, Germany), 5 µl 10x MOPS buffer (200 mM MOPS; 50 mM Sodium acetate; 10 mM EDTA; Carl Roth GmbH, Karlsruhe, Germany) and 10 µl 37% formaldehyde (Carl Roth GmbH, Karlsruhe, Germany), and incubated 5 min at 65°C for denaturation, followed by adding 2 µl ethidium bromide (10 mg/ml; Carl Roth GmbH, Karlsruhe, Germany) to the mix and electrophoresis in 1x MOPS buffer at 85 V for 2.5 hours. RNA was visualized in the gel by excitation under UV light using the BioDocAnalyze Gel-analyze System (Biometra GmbH, Göttingen, Germany). Formaldehyde is toxic through skin contact and inhalation of vapors thus all manipulations involving formaldehyde had to be performed in a chemical fume hood.

2.8 cDNA synthesis and reverse transcription PCR (RT-PCR)

Before messenger RNA (mRNA) could be transcribed into complementary DNA (cDNA) by reverse transcription, genomic DNA had to be removed from RNA preparations by incubation with RNase-free DNase I (Fermentas, St. Leon-Rot, Germany) according to the manufacturer's instructions. The reactions were carried out in a total volume of 10 µl containing 1 µg of RNA, 1 µl of 10x reaction buffer with $MgCl_2$ (100 mM Tris-HCl, pH=7.5; 25 mM $MgCl_2$; 1 mM $CaCl_2$) and 1 unit of DNase I (1 U/µl). Samples were incubated at 37°C for 30 min, followed by adding 1 µl of 50 mM EDTA and further incubation for 10 min at 65°C. Incubation was performed on a thermal cycler (Life Technologies GmbH, Darmstadt, Germany) with lid temperature at 104°C. The DNase I treated RNA was then used as a template for cDNA synthesis. Reverse transcription was performed using the iScript™ cDNA Synthesis Kit (Bio-Rad Laboratories GmbH, München, Germany) following the manufacturer's protocol. The DNase-I treated RNA was used as template in a total volume of 20 µl containing 4 µl of 5x iScript reaction mix (10 mM dNTPs; 0.5 mM oligo dT; 50 mM $MgCl_2$; and stabilizers), 1 µl of iScript reverse transcriptase and nuclease-free water. The reaction mix was incubated in a thermal cycler (Life Technologies GmbH, Darmstadt, Germany) with lid temperature at 104°C during the following cycling profile: 25°C for 5 min, 42°C for 45°C, 85°C for 5 min and finally hold at 4°C. The obtained cDNA was subsequently used as template in downstream RT-PCR analysis.

RT-PCR was set up in a total volume of 20 µl containing 2 µl of cDNA template, 2 µl of 10x PCR buffer (100 mM Tris-HCl, pH 8.3; 500 mM KCl; 15 mM $MgCl_2$; 0.01% gelatin), 2 µl of dNTPs (40 mM), 1 µl each of 5 µM forward and reverse primer, 0.1 µl of HotStarTaq DNA polymerase (5 U/µl; Qiagen, Düsseldorf, Germany) and 11.9 µl of nuclease-free water. A touch-down PCR program was used as mentioned above in the section Polymerase Chain Reaction (Section 2.4).

2.9 PCR product purification and Sanger sequencing

The PCR product was purified using the NucleoFast® 96 PCR Kit (Macherey-Nagel, Düren, Germany). Ultrafiltration was performed on a NucleoVac 96 vacuum manifold (Macherey-Nagel, Düren, Germany). Purification procedures were as follows: The PCR product mixture was filled up to 100 µl with nuclease-free water and transferred

to the NucleoFast® filter plate, vacuum until no droplet formed. Repeat the vacuum procedure after adding additional 100 µl of nuclease-free water. The purified PCR product was dissolved in 30 µl nuclease-free water, and transferred into new 96-well plate (Fisher Scientific GmbH, Schwerte, Germany) after shaking 10 min on Titramax 100 (Heidolph Instruments GmbH, Schwabach, Germany). The concentration of purified PCR products was determined visually by agarose gel electrophoresis along with loading a defined dilution series (1 µg – 100 ng) of λDNA (Fermentas GmbH, St. Leon-Rot, Germany). Sequence data was generated by cycle sequencing with BigDye Terminator (BigDye®Terminator v3.1, Applied Biosystems, Darmstadt, Germany) chemistry using PCR products as template. Samples sent for sequencing were prepared with a total volume of 6 µl containing 5 µl of purified PCR product with normalized concentration (ca. 10 ng/100 bp) plus 1 µl of forward/reverse primer (5 µM each) as used for initial PCR.

Sanger cycle sequencing was carried out in-house at the Plant Genome Resources Center (PGRC) of Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. Sample preparation was carried out according to the sequencing protocol [5 µl of PCR product (10ng/100bp) plus 1 µl of sequencing primer (5 µM)] and sequenced by using Big Dye Terminator chemistry and an ABI 3730 XL instrument (Life Technologies GmbH, Darmstadt, Germany).

2.10 Sequence analysis

DNA sequences were obtained by PHRED base-calling and bases with a PHRED score of 20 and above were remained after quality-trimming (Ewing and Green, 1998; Ewing et al., 1998). The quality-trimmed sequences were used for further analysis, e.g. SNP calling for marker development, sequence alignment to confirm the cloning product, using the Sequencher® version 5.2.3 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI USA. http://www.genecodes.com).

2.11 Marker development

2.11.1 Primer design

The primers used for marker development were designed using the online software Primer3 v. 0.4.0 (http://bioinfo.ut.ee/primer3-0.4.0/) (Koressaar and Remm, 2007; Untergasser et al., 2012). Default parameters were used with minor modifications:

GC content was set within the range of 50% - 60% and the product size to a range between 300 and 1700bp (Appendix Table 3). Alternatively, primers used for amplification of the fragments, i.e. *albostrians* (*as*) gene, promoter of *as* gene, CCT domain and predicted chloroplast transit peptides (cTP) of *as* gene, were selected manually according to the 5' and 3' sequences of the respective cloning fragment and proper restriction sites were introduced to both ends of the amplified fragment to facilitate cloning of the fragments. The primers used for transcription activator-like effector nuclease (TALEN) construction and targeting induced local lesions in genomes (TILLING) experiments were also selected manually. Two criteria were followed for manually assign primers: GC content with a range of 50-60% and DNA melting temperature (Tm) of around 60°C.

2.11.2 Cleaved Amplified Polymorphic Sequences (CAPS) marker development and analysis

After identifying single nucleotide polymorphisms (SNPs) between wild type (WT) and mutant (MT) parental genotypes through sequence analysis as outlined above (section 2.10), sequences were analyzed by help of the SNP2CAPS software (Thiel et al., 2004), for the selection of appropriate restriction enzymes (common enzyme, low price, high specificity under uniform standard incubation conditions) that would allow to clearly differentiate WT and MT alleles after restriction analysis of PCR products based on fragment size polymorphism. DNA digestion was performed in a 10 µl volume containing 5 µl of PCR product, 1 µl of proper 10x buffer (New England Biolabs, Hitchin, UK), 1 unit of enzyme (New England Biolabs, Hitchin, UK) and adjusted to final volume by adding ddH_2O. The reaction mix was incubated for one hour in an oven with constant temperature (37°C).

For the design of CAPS markers numerous genomic resources could be utilized, including a high-density transcript linkage map (Sato et al., 2009a), virtually ordered gene map of barley (Genome Zipper, Mayer et al., 2011), and the physical map of barley (International Barley Genome Sequencing Consortium, 2012).

2.12 GoldenGate® genotyping assay with 384-plex veraCode® technology

GoldenGate® genotyping assay combined with custom 48-, 96-, 144-,192, and 384-plex veraCode® technology (Illumina, Inc., San Diego, CA, USA) provides one of the most robust and flexible platforms for SNP genotyping. As of to date, it has been

widely applied in plants and animals, such as pea (Deulvot et al., 2010), peanuts (Bertioli et al., 2014), rice (Chen et al., 2014), and pig (Zhang et al., 2014), for linkage and association genetic studies. In this study, for genotyping the bi-parental populations MM4205 and BM4205 we took advantage of an Illumina GoldenGate® genotyping assay composed of a pre-selected set of 381 informative barley oligonucleotide pool assay (BOPA) SNPs (Close et al., 2009) that were evenly distributed across the barley genome (Appendix Table 4). This initial genotyping of the basic population of 91 F_2 individuals was provided as a service in collaboration by KWS, Einbeck (kindly provided by Dr. V. Korzun). The genotyping calls were obtained by help of the GenomeStudio software (Illumina, Inc., San Diego, CA, USA) under default parameters.

2.13 Genetic linkage analysis

Genetic linkage analysis was performed using Joinmap4.0 software (Van Ooijen, 2006) as described by the manual's instructions. Homozygous WT, heterozygous and homozygous mutant allele calls were defined as A, H and B, respectively; missing data was indicated by a dash. Maximum Likelihood algorithm and Kosambi's mapping function were chosen for building the linkage maps. Markers were grouped into seven groups based on Logarithm of Odds (LOD = 4) groupings. Visualization of maps derived from Joinmap4.0 was achieved by MapChart software (Voorrips, 2002).

2.14 PCR screening of a barley BAC library

In order to integrate newly developed markers with the physical map of barley, a PCR-based physical map anchoring approach was used to screen multi-dimensional DNA pools of a barley BAC library. Construction of the barley BAC library HVVMRXALLeA was reported by Schulte et al. (2011). Multi-dimensional DNA pools of 147,840 BACs were provided based on proprietary Matrix pooling and superpooling system of the company Amplicon Express (WA, USA). Screening of the resource was following the manufacturer's protocols and was described earlier (Ariyadasa and Stein, 2012; International Barley Genome Sequencing Consortium, 2012). In brief, the 147,840 BACs were distributed over three hundred and eighty-five 384-well plates. All of the 385 plates were further assigned into 55 superpools, each superpool containing seven consecutive 384-well plates. Subsequently, clones of

each of the seven plates, of the respective 16 plate rows and of the respective 24 plate columns across all 7 plates of a given superpool were pooled, to form 7 plate pools, 16 row pools and 24 column pools per set of seven plates. The created plate, row and column pools from each superpool were further combined to create the three dimensional matrix pools. Finally, screening was performed in 5 matrix plate pools, 8 matrix row pools and 10 matrix column pools. The detailed illustration of the superpool and matrix pool design can be found under http://ampliconexpress.com/.

2.15 Annotation of DNA and protein sequence and domains of the *albostrians* gene

Chloroplast transit peptides of the deduced ALBOSTRIANS (AS) protein was predicted by SignalP 4.1 (Petersen et al., 2011), and *in silico* prediction of subcellular localization of the AS protein was performed through ChloroP 1.1 (Emanuelsson et al., 1999), TargetP 1.1 (Nielsen et al., 1997; Emanuelsson et al., 2000), Predator 1.03 (Small et al., 2004), PredSL (Petsalaki et al., 2006) and WoLF PSORT (Horton et al., 2007). The functional domains of AS protein were determined by InterProScan 5 (Jones et al., 2014).

2.16 Functional validation of the gene *albostrians* by TILLING

TILLING (Targeting Induced Local Lesions IN Genomes) is a reverse genetics method that allows directed identification of mutations in a specific genomic region. The principle of TILLING is based on the formation of heteroduplex DNA at mutated loci if PCR amplicons are generated from a pool of WT and MT DNA. Non-paired DNA strands of such heteroduplex DNA will remain non-paired at SNP positions. These single stranded regions can be cleaved by a plant endonuclease *Cel* I (Transgenomics, Omaha, USA) (Oleykowski et al., 1998). In this study, an ethyl methanesulfonate (EMS) induced TILLING population, comprising 10,279 M_2 plants, derived from a two-rowed malting barley cultivar ‘Barke’ (Gottwald et al., 2009), was used for identification of independent mutated alleles of the *albostrians* gene. In brief, two pairs of primer were designed covering the coding sequence of the candidate *albostrians* gene. PCR amplicons were analyzed combined with dsDNA Cleavage Kit (DNF-480-3000) and Gel-dsDNA reagent kit (DNF-910-K1000) according to the manufacturer’s protocols (Advanced Analytical Technologies GmbH, Heidelberg, Germany). Subsequently, the cleaved PCR products were separated using the

*Advan*CE™ FS96 capillary electrophoresis system (Advanced Analytical Technologies GmbH, Heidelberg, Germany) and results were interpreted with assistance of the PRO Size™ software (Advanced Analytical Technologies GmbH, Heidelberg, Germany).

The identified M_2 TILLING mutants were confirmed by Sanger sequencing of PCR amplicons derived from the respective families. Plant families carrying non-synonymous mutations, deletions, or immature stop codons were selected for phenotyping and multiplication. Phenotyping was performed in M_3 and M_4 generation of the identified M_2 mutants. Heterozygous plants were propagated and maintained for further reproduction.

2.17 Vector construction for functional validation and subcellular localization

2.17.1 Complementation of mutant plants by overexpression of a WT gene copy

Wild type and mutant versions of the *albostrians* gene were obtained by RT-PCR using cDNA of 'Haisa' and 'M4205' as template, respectively. Fragments amplified from 'M4205' carrying a 4 bp deletion at position 1123 - 1126 bp (count from adenine of start codon as +1) compared to the wild type. The insert was introduced into the *Hinc*II cloning site of pUbi-ABM (DNA Cloning Service, Hamburg, Germany) in a reaction volume of 20 µl containing 0.4 µl of digested pUbi-ABM (100 ng), 2 µl of RT-PCR product, 2 µl of 10x T4 DNA ligase buffer (100 mM $MgCl_2$; 100 mM DTT; 5 mM ATP; 400 mM Tris-HCl; pH 7.8 at 25°C), 2 µl of 50% PEG 4000 solution, 1 µl of T4 DNA ligase (Fermentas GmbH, Schwerte, Germany) and 12.6 µl nuclease-free water. The reaction mix was incubated in a thermal cycler (Life Technologies GmbH, Darmstadt, Germany) with the program 22°C for 2 hours, followed by 70°C for 10 min to inactivate the enzyme. Thereafter, the ligated reaction mix was further digested at 37°C for 2 hours after adding 3 µl of 10x buffer 3.1 (New England Biolabs GmbH, Frankfurt am Main, Germany), 1 unit of *Hinc*II (New England Biolabs GmbH, Frankfurt am Main, Germany) and 6 µl nuclease-free water (*Hinc*II generates blunt end, this step aims to digest the religation vectors with no desired insert). Two microliter of the reaction mix was used for transformation of TOP10 competent cells (Life Technologies GmbH, Darmstadt, Germany) by the heat shock method described elsewhere (Froger and Hall, 2007). Positive clones were checked via restriction analysis and confirmed by Sanger sequencing. Subsequently, the

expression cassette of the WT/MT *albostrians* gene was introduced into the *Sfi*I cloning site of the binary vector p6d35S (Nagy et al., 2011). Binary vectors containing desired insert were transformed into *A. tumefaciens* strain AGL1 via electroporation (SAMBROOK et al., 1989). *Agrobacterium*-mediated transformation was performed by Dr. Götz Hensel in the research group of Plant Reproductive Biology (AG PRB) following the protocol described elsewhere (Himmelbach et al., 2007; Hensel et al., 2008). The derived constructs information is summarized in Table 2-2.

2.17.2 Inducing knock out (KO) mutant lines by a TALEN approach

TALEN (Transcription Activator-Like Effectors Nuclease) is an artificially engineered nuclease with the potential of genome editing as a result of causing double strand breaks (DSB), and its principle and application had been reviewed elsewhere (Joung and Sander, 2013; Puchta and Fauser, 2013; Voytas, 2013). The first report in barley using TALEN was to knock out the *phytase* gene (Wendt et al., 2013). Recently, expression of the functional *GFP* gene integrated in a transgenic barley line was successfully silenced through transformation with a *GFP*-specific TALEN pair (Gurushidze et al., 2014).

The module, array and last repeat TALEN plasmids ordered from Addgene (http://www.addgene.org/) were assembled in the research group of PRB as guided by the Golden Gate TALENs assembly protocol (Cermak et al., 2011). The assembled forward TALEN arm (pTAL_*HvAs*_F) and reverse TALEN arm (pTAL_*HvAs*_R) were each introduced into the *Asc*I/*BamH*I and *Spe*I/ *BamH*I cloning sites of pSH60 and pSH34 (optimized backbone plasmids generated by Stefan Hiekel, AG PRB, IPK), respectively. The “self-cleaving” 2A peptides (T2A) were used to generate multiple proteins from a single promoter in many applications (de Felipe, 2004; Osborn et al., 2005; de Felipe et al., 2006). Hence, TALEN units of the two intermediate vectors were subsequently cloned into a 2A peptide-enabled dual expression vector pSH68 via *Asc*I/*Bcl*I (*Bcl*I produces the same sticky overhangs as *BamH*I) and *Spe*I/*BamH*I digestion. The TALEN expression cassette (pSH68_TALEN_*HvAs*_F_T2A_TALEN_*HvAs*_R) was finally introduced into the *Sfi*I cloning site of the binary vector p6id35STE9, which harbors *hpt*, a gene which confers hygromycin resistance, driven by the cauliflower mosaic virus double enhanced 35S (CaMVd35S) promoter. The binary vector containing both TALENs

(p6id35STE9_TALEN_*HvAs*_F_T2A_TALEN_*HvAs*_R, with an alternative name pSH83 for convenient labeling) was then introduced into *A. tumefaciens* strain LBA4404pSB1 through electroporation (SAMBROOK et al., 1989). Aliquots of the liquid culture containing 15% glycerol were stored in -70°C for the downstream *Agrobacterium*-mediated transformation.

2.17.3 Subcellular localization of the WT and MT ALBOSTRIANS proteins

The coding sequence of the WT *albostrians* gene of barley cv. 'Haisa', as well as its two MT alternatives from M4205 and M_2-TILLING mutant 6460-1 were fused to the N-terminus of the GFP reporter gene (Chiu et al., 1996), respectively, through ligation into the *Spe*I/*Hind*III cloning sites of the vector pSB179 (provided by Dr. Jochen Kumlehn), which contained an ampicillin resistance gene as selection marker. The digestion reactions were performed in a total volume of 10 µl containing 1 µg of plasmid, 5 µl of insert fragment, 1 µl of FastDigest *Spe*I (Fisher Scientific GmbH, Schwerte, Germany), 1 µl of FastDigest *Hind*III (Fisher Scientific GmbH, Schwerte, Germany), 1 µl of 10x FastDigest green buffer, and nuclease-free water. After incubation at 37°C for 20 min, the digested products were separated on a 2% (w/v) agarose gel and the expected fragments were recovered and purified using MiniElute Gel Extraction Kit (QIAGEN GmbH, Hilden, Germany). Ligation reactions were carried out in a total volume of 20 µl after mixing 1 unit of T4 DNA ligase (Fermentas GmbH, Schwerte, Germany), 2 µl of 10x ligation buffer (100 mM $MgCl_2$; 100 mM DTT; 5 mM ATP; 100 mM Tris-HCl, pH=7.8), 3 µl of purified PCR product, 1 µl of purified backbone fragment, and nuclease-free water. Reaction mix was incubated at 22°C for 2 hours. Then, the recombinant plasmid was transformed into TOP10 competent cell (Life Technologies GmbH, Darmstadt, Germany) via a heat shock method as described by Froger and Hall (2007). The obtained recombinant expression cassette was confirmed via Sanger sequencing and applied for the downstream bombardment experiment. The derived constructs information is summarized in Table 2-2.

Table 2-2: Summary of the constructed vectors.

Construct ID	Source (Organism)[A]	Genotype Background[B]	Application
p6d35S-*HvAs*_WT	Barley	Haisa	Complementation
p6d35S-*HvAs*_M4205	Barley	Haisa	Complementation
pSB179-*HvAs*_WT	Barley	Haisa	Subcellular localization
pSB179-*HvAs*_M4205	Barley	Haisa	Subcellular localization
pSB179-*HvAs*_TILLING	Barley	Haisa	Subcellular localization
pSH83	-	-	Inducing KO mutant lines

[A]– The origin of the insertion fragment.

[B]– The barley genotype used for amplification of the insertion fragment.

2.18 Biolistic transient expression assay for subcellular localization of the ALBOSTRIANS protein

The principle of particle bombardment/biolistic delivery is that microscopic DNA-coated gold particles are accelerated at high speed by helium gas with a vacuum and travel at such a velocity as to penetrate target cells or intact tissues (Sparks and Jones, 2009). In the present study, the bombardment experiment was performed by using the Biolistic™ Particle Delivery System PDS-1000 (Bio-Rad Laboratories GmbH, München, Germany). The physical parameters, 1100 psi acceleration pressure and 27 inch Hg vacuum pressure, had been set to reach efficient bombardment conditions for barley epidermal cells. Preparation and delivery of DNA-coated gold particles was performed according to Sparks and Jones (2009). In brief, thirty microgram of gold powder (particle diameter of 1.0 µm; Bio-Rad Laboratories GmbH, München, Germany) was dissolved in 1 ml of 100% ethanol under vortex. Gold suspension was precipitated through centrifugation (13,000 x *g*, 1 min, room temperature). Discard the supernatant and re-suspend the gold in 1 ml of 100% ethanol for subsequent DNA coating or store at -20°C. Fifty microliter of stock gold suspension solution (30 mg/ml) was transferred into the Eppendorf® Biopur® Safe-Lock tube (Sigma-Aldrich Chemie GmbH, Munich, Germany) and washed with 100 µl of nuclease-free water. Gold particles were precipitated by centrifugation (13,000 x *g* for 2 min at 20°C) and recovered after removal of the supernatant. This washing step was repeated three times. Subsequently, the gold pellet was dissolved in 25 µl of nuclease-free water followed by sonication for 1 min. Thereafter, 5 µl of plasmids, which containing 2.5 µl of fusion construct (1 µg/µl) and 2.5 µl of plastid organelle marker pt-rk-CD3-999 (1 µg/µl) (Nelson et al., 2007), 10 µl of 0.1 M spermidine (Sigma-Aldrich Chemie GmbH, Munich, Germany) and 25 µl of 2.5 M $CaCl_2$ (Carl Roth GmbH, Karlsruhe, Germany) were added and the final mixture vortexed for 2

min. DNA-gold-pellet was collected by centrifugation (13,000 x *g* for 2 min at 20°C). After removal of the supernatant, the DNA-gold-pellet was washed twice with 80 µl of 100% ethanol. The obtained DNA-gold-pellet was re-suspended in 60 µl of 100% ethanol. Five microliter of the DNA-coated microparticles suspension was loaded on the center of a macrocarrier (Bio-Rad Laboratories GmbH, München, Germany), desiccated and used for bombardment.

Plants of barley cv. 'Haisa' were grown under greenhouse conditions (16 h light / 20°C and 8 h dark / 15°C cycle). First leaves of the seedlings at 2-leaf stage were collected and placed on medium, containing 1% (w/v) of phytoagar (Duchefa Biochemie, Haarlem, Netherlands), 20 µg/ml of Benzimidazol and 10 µg/ml of chloramphenicol, for 4 h prior and 24 h after bombardment.

The sample (around 1 cm^2 leaf area) was immersed in a pre-treated 2 ml Eppendorf tube (with holes on its cap) containing water. Then, a vacuum was applied until no bubbles released from the stomata cells. The sample was subsequently transferred to a slide using forceps and covered with a coverslip followed by sealing with Elmer's white glue (Elmer's Products, Inc., Ohio, USA). The sample was then scanned for the presence of fluorescence signals by help of a Confocal Laser Scanning Microscopy LSM 780 (Carl Zeiss MicroImaging GmbH, Jena, Germany).

2.19 Phylogenetic analysis

The ALBOSTRIANS protein sequence was used as BLASTp (Mount, 2007) query to retrieve homologs from other species on NCBI and phytozome (Goodstein et al., 2012) databases. Phylogenetic analysis was performed using MEGA6 (Tamura et al., 2013) following the protocol of Hall (2013). The evolutionary history was inferred by using the Maximum Likelihood (ML) method based on the JTT matrix-based model. The phylogenetic tree was constructed using ML heuristic method Nearest Neighbor Interchange (NNI) implemented in MEGA6 (Tamura et al., 2013). Initial trees for heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using a JTT model. A discrete Gamma distribution was used to model evolutionary rate differences among sites [5 categories (+G, parameter = 0.8412)]. The bootstrap consensus tree inferred from 1,000 replicates was taken to represent the evolutionary history of the analyzed sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5%

alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 215 positions in the final dataset.

3 Results

3.1 Genetic mapping of the barley gene *albostrians*

3.1.1 Low-resolution genetic mapping

Over the last 30 years, studies involving the *albostrians* mutant mainly focused on plastid development and plastid-nucleus interactions. The genetic position of the *albostrians* locus, however, had never been determined. Hence, with the aim to allocate the genetic position of the gene *albostrians*, a first genetic mapping was initiated. Two independent F_2 populations, Barke x M4205 (BM4205) and Morex x M4205 (MM4205) were constructed. For the initial allocation of the gene a set of 91 genotypes of each population was screened by an Illumina Golden Gate assay based on a set of SNP markers (Appendix Table 4) that were selected for their even genome distribution and balanced allele frequency in previously analyzed modern European barley germplasm (personnel communication Dr. Viktor Korzun). An overall (global) genotyping success rate of 98.16% (374 SNPs) and 97.90% (373 SNPs) of the total 381 SNP markers was obtained for BM4205 and MM4205 mapping populations, respectively. Out of the successfully genotyped loci, 130 SNPs (34.76%) were found to be polymorphic in population BM4205 and 198 SNPs (53.08%) were polymorphic in population MM4205 (Table 3-1). After scoring the leaf phenotype (fully green, variegated, *albino*) of all F_2 individuals and 30 plants of their respective F_3 progeny of both populations (Appendix Table 2), a genetic map was calculated and the gene *albostrians* could be reliably allocated to the long arm of chromosome 7H in a 14.29 cM genetic interval between two flanking markers 2_0771 and 1_0169 in the MM4205 mapping population (Figure 3-1).

Table 3-1: Summary of the SNPs (GoldenGate assay) used in BM4205 and MM4205 mapping populations

	BM4205	MM4205
Total	381	381
Successful (Feasible SNPs)	374 (98.16%)	373 (97.90%)
Failures	7 (1.84%)	8 (2.10%)
Monomorphic	244 (65.24%)	175 (46.92%)
Polymorphic	130 (34.76%)	198 (53.08%)

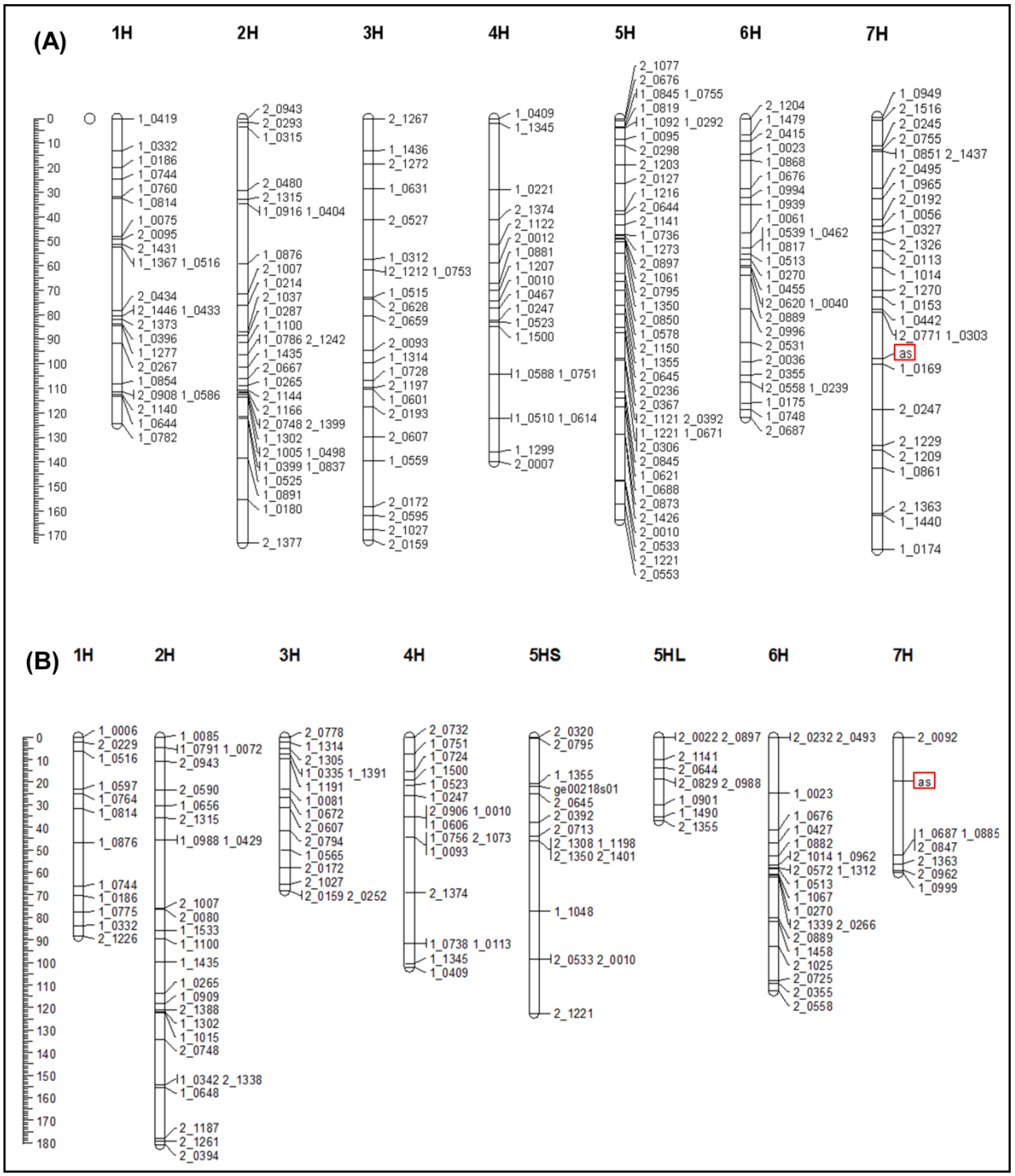

Figure 3-1: Low-resolution mapping of the *albostrians* gene. Low-resolution genetic mapping of the gene *albostrians* in the MM4205 and BM4205 mapping populations with 91 F_2 genotypes, respectively, based on a 384-plex Veracode SNP genotyping assay. (A) Seven linkage groups are derived from 198 polymorphic markers within the MM4205 population (LOD = 4.0). (B) In BM4205 population, eight linkage groups are generated, with markers of chromosome 5H being separated into two linkage groups (LOD =4.0). In both populations, the *albostrians* gene could be allocated to the genetic position on chromosome 7H indicated by the red box.

3.1.2 Marker saturation for fine mapping of the gene *albostrians*

After obtaining an initial mapping position of the gene *albostrians,* new markers were required for high-resolution mapping and cloning of the gene. The parental genotypes of mapping population BM4205 revealed a lower level of polymorphism in the initial genetic mapping experiments which indicated that polymorphism survey for new marker development would most likely be less efficient in this population. Therefore, priority for further mapping of the gene *albostrians* including the step of new marker development and saturation of the target interval was given to population MM4205.

Marker development relied on publicly available genomic resources and a newly generated resource of survey sequencing of the mutant line in frame of this thesis. During the past decade, several barley genetic maps were developed based on several different bi-parental or DH mapping populations (e.g. Rostoks et al., 2005; Close et al., 2009; Sato et al., 2009a). With this information, it was possible to directly select markers from the published genetic maps that were potentially located within the target interval and apply them to mapping population MM4205 with first testing polymorphism on the parental genotypes. A transcript linkage map of barley (Sato et al., 2009a) was exploited first for marker development. As the initial step, anchoring the flanking markers 2_0771 and 1_0169 to the original genetic map (Close et al., 2009) revealed a genetic interval from positon 87.21 to 104.78 cM. Subsequently, markers within this target interval on the transcript map (Sato et al., 2009a) were selected as the initial step of saturation mapping the gene *albostrians*. In order to further narrow down the target interval, the newly identified distal flanking marker CAPS_2536 and the proximal flanking marker 1_0169 were anchored to the barley 'Genome Zipper', which represented a high density linear gene order of barley based on the integration of a high density genetic map with conserved synteny information of related and sequenced grass genomes (Mayer et al., 2011). Then, the sequence of barley full length cDNAs (FLcDNAs) that were predicted to be located within the target interval were used as a template for gene-based marker development (Matsumoto et al., 2011; Mayer et al., 2011). In duration of the project the sequence-enriched physical map of barley (International Barley Genome Sequencing Consortium, 2012) became available and the associated whole genome shotgun (WGS) sequence information provided a new genomic resource for marker

development. Similar to the strategy for developing the gene-based markers, after anchoring the flanking markers Zip_2613 and 1_0169 to the physical map, the WGS sequences within the target interval were selected as genomic resource for new marker design.

In parallel with marker development by taking advantage of the publicly available resources, whole genome survey sequencing of the mutant line M4205 as well as cultivar 'Haisa' was initiated in order to accelerate identification of sequence polymorphisms for efficient marker design and mapping of the *albostrians* gene. The WGS data was mapped to the Morex whole genome shotgun reference sequence assembly (WGS_Morex_assembly; International Barley Genome Sequencing Consortium, 2012) followed by SNP mining. As a result, 112,989 SNPs were identified between Morex and M4205. This information greatly facilitated the direct design of new markers for saturation mapping of the target interval.

Of the 42 markers selected from a published transcript map (Sato et al., 2009a), 21 markers were polymorphic, however, five marker polymorphisms could not be reproduced in the MM4205 population. The remaining 16 polymorphic markers were further divided into 7 bins based on their genetic positions, as a result, one marker from each bin was selected to saturate the genetic interval. In addition, 15 out of 43 FLcDNA-derived amplicons could be converted into polymorphic markers and ten non-redundant markers of this group could be mapped to the target interval. Moreover, 55 specific amplicons were obtained by taking advantage of the WGS sequence information of the two parental genotypes of the mapping population resulting in 27 polymorphic markers; eight of them finally used for marker saturation of the target interval. Overall, 25 new markers could be mapped to saturate the target interval for mapping the gene *albostrians* (Table 3-2, Table 3-3, Appendix Table 3).

With the 91 genotypes of the MM4205 mapping population, saturation mapping of the gene *albostrians* narrowed down the original 14.29 cM target interval to 6.04 cM between flanking markers Contig_220966 and 1_0169. Besides, a cluster of co-segregating markers was identified within the target region, which provided effective information for high-resolution genetic mapping as to identify recombination events within a larger population (Figure 3-2A). Notably, in addition to saturation mapping of the genetic interval, segmental recombination inbred lines (RILs) were selected from

30 F_3 progeny plants of each of the 14 recombinant F_2 plants (between marker CAPS_2536 and CAPS_2560). And totally 13 F_3 segmental RILs were selected from the 14 recombinant plant families.

Table 3-2: Summary of the markers employed for saturation mapping.

Marker Resource	Marker Type	SNPs	Indels	Cumulative Size of the Amplicon (bp)	SNP Frequency	INDEL Frequency	# of primer pairs	Specific Amplicon Obtained	# of Polymorphic Markers	# of Markers Used for Saturation Mapping
Sato et al. (2009a)	CAPS	40	4	24,312	1/608 bp	1/6,078 bp	42	42	21 (50.00%)	7
Mayer et al. (2011)	CAPS	36	6	34,556	1/957 bp	1/5,743 bp	45	43	15 (34.88%)	10
IBSC (2012)	CAPS	102	14	52,786	1/518 bp	1/3,770 bp	60	55	27 (49.09%)	8
Total (average)		178	24	111,654	1/627 bp	1/4,652 bp	147	140	63 (45.00%)	25

Table 3-3: Summary of polymorphic markers.

Order[1]	Primer ID	Genetic position (cM)[2]	SNPs[3]	Del[4]	Marker type	Enzyme[5]	Morex (bp)[6]	M4205 (bp)	Note[7]
3	CAPS_2503 (k06406)	77.5495750708	1	-	CAPS	*ScrF*I	411, 379	790	Mapped
18	CAPS_2536 (k08949)	85.9773371105	7	-	CAPS	*SfaN*I	552,182	734	Mapped
26	CAPS_2551 (k03358)	92.0679886686	2	-	CAPS	*Ban*I	444	232,212	Mapped
30	CAPS_2558 (k00397)	97.2733711048	2	-	CAPS	*BstU*I	347,121	468	Mapped
31	CAPS_2560 (k09456)	95.5028328612	3	-	CAPS	*Sau96*I	355,318	673	Mapped
33	CAPS_2562 (k06299)	102.195467422	2	1	CAPS	*Nsi*I	371	197,175	Mapped
42	CAPS_2586 (k02011)	107.577903683	2	-	CAPS	*Taq*I	195,95,60	195,155	Mapped
45	3_0168	91.9263456091	3	3	CAPS	*Taq*I	559,224	389,383	Mapped
53	Zip_2601	103.541076487	4	-	CAPS	*MspA1*I	592,218	366,226,218	Mapped
56	Zip_2613	85.9773371105	2	1	CAPS	*AlwN*I	625,229	853	Mapped
66	Zip_2656	91.7847025495	1	-	CAPS	*Cac8*I	436,240,121	347, 240, 121, 89	Mapped
67	Zip_2661	91.9263456091	1	-	CAPS	*Acc*I	843	660,183	Mapped
68	Zip_2662	92.209631728	4	1	CAPS	*Mme*I	699	569,126	Mapped
69	Zip_2665	92.0679886685	2	-	CAPS	*AlwN*I	959	722,237	Mapped
73	Zip_2667_4	91.0552407932	2	-	CAPS	*EcoR*I	692	380,312	Mapped
78	Zip_2672	92.0679886686	3	-	CAPS	*Fau*I	613	497, 116	Mapped
79	Zip_2680_1	92.209631728	2	-	CAPS	*Sau96*I	777,73	460,317,73	Mapped

88	Contig_220966	86.2709513692	6	-	CAPS	*Acc*I	518,84,67	602,67	Mapped
93	Contig_1596897	86.6540761725	6	-	CAPS	*HinP1*I	405,278	683	Mapped
117	Contig_37952_1	91.9263456091	2	-	CAPS	*BstY*I	485,365	365,297,188	Mapped
123	Contig_40728_2	92.0679886686	2	2	CAPS	*HinP1*I	774,187	950	Mapped
125	Contig_49785_5	92.0679886686	6	-	CAPS	*Hpa*II	757,33	520,237,33	Mapped
142	Contig_1561286_2	92.0679886686	7	-	CAPS	*HinP1*I	788,300,205	993,3	Mapped
146	Contig_1575446	92.0679886686	7	-	CAPS	*Nsp*I	618,18	798	Mapped
148	Contig_92279	91.9263456090	11	1	CAPS	*Acc*I	766,76	425,340,76	Mapped
1	CAPS_2471 (k08084)	74.2917847025	3	-	CAPS		Not suitable for CAPS assay		Not mapped
4	CAPS_2504 (k07980)	77.2662889518	1	-	CAPS		Not suitable for CAPS assay		Not mapped
5	CAPS_2505 (k02916)	-	1	-	CAPS	*Ssp*I	1076	758,318	Not mapped
6	CAPS_2506 (k02987)	76.6997167139	1	-	CAPS		Not suitable for CAPS assay		Not mapped
7	CAPS_2507 (k02572)	77.4079320113	2	-	CAPS		Not suitable for CAPS assay		Not mapped
20	CAPS_2541 (k04151)	85.9773371105	1	-	CAPS		Not suitable for CAPS assay		Not mapped
23	CAPS_2547 (k09929)	91.428470255	2	-	CAPS	*Mme*I	260,88,6	165,101,88	Not mapped
24	CAPS_2548 (k04467)	91.9263456091	2	1	CAPS	*Apo*I	194,94,61	286, 61	Not mapped
25	CAPS_2549 (k03809)	92.0679886685	1	-	CAPS	*Set*I	204,95,60,22	264,95,22	Not mapped
27	CAPS_2552 (k03428)	91.9263456091	2	-	CAPS	*Eci*I	419	294,125	Not mapped
29	CAPS_2557 (k03102)	91.9263456091	1	1	CAPS	*Rsa*I	465,286,70	533,286	Not mapped
36	CAPS_2565 (k08205)	100	1	-	CAPS	*Pen*I	132,103,101,42	204,132,42	Not mapped
37	CAPS_2568 (k07247)	100	2	-	CAPS	*Acl*I	1608	1462, 146	Not mapped
39	CAPS_2574 (k03027)	100	3	1	CAPS	*Cac8*I	528	302,223,4	Not mapped
70	Zip_2667_1	91.0552407932	5	1	CAPS	*Esp3*I	954	803,149	Not mapped
72	Zip_2667_3	91.0552407932	1	-	CAPS		Not suitable for CAPS assay		Not mapped
76	Zip_2671_1	91.9971671389	2	-	CAPS	*Hin4*II	970	826,144	Not mapped
77	Zip_2671_2	91.9971671389	1	-	CAPS		Not suitable for CAPS assay		Not mapped
80	Zip_2680_2	92.209631728	3	-	CAPS	*Sau96*I	264,225,80,76	489,80,76	Not mapped
92	Contig_1586971	86.6324362606	2	1	CAPS	*Bbv*I	661	532, 128	Not mapped
99	Contig_39427	88.9134796978	-	2	CAPS		Not suitable for CAPS assay		Not mapped

102	Contig_157273	99.0793201133	-	1	CAPS		Not suitable for CAPS assay		Not mapped
103	Contig_72015	89.6338526912	2	-	CAPS		Not suitable for CAPS assay		Not mapped
118	Contig_37952_2	91.9263456091	4	1	CAPS	*AlwF*I	922	545,376	Not mapped
119	Contig_37952_3	91.9263456091	1	2	CAPS		Not suitable for CAPS assay		Not mapped
120	Contig_37952_4	91.9263456091	2	1	CAPS	*Nci*I	986	562,421	Not mapped
121	Contig_37952_5	91.9263456091	7	2	CAPS	*Eco105*I	559, 290	839	Not mapped
122	Contig_40728_1	92.0679886686	3	-	CAPS	*SfaN*I	745,132,99	721,132,99,24	Not mapped
124	Contig_40728_3	92.0679886686	7	-	CAPS	*Bfa*I	501,447,224	671,501	Not mapped
126	Contig_49785_2	92.0679886686	2	1	CAPS		Not suitable for CAPS assay		Not mapped
127	Contig_49785_3	92.0679886686	4	-	CAPS	*Hinf*I	898,298	478,420,298	Not mapped
128	Contig_49785_4	92.0679886686	8	-	CAPS	*Bbv*I	990	624,366	Not mapped
132	Contig_137310_2	91.9263456091	9	-	CAPS	*Fok*I	726,454	1180	Not mapped
134	Contig_275079	92.0679886686	1	-	CAPS		Not suitable for CAPS assay		Not mapped
135	Contig_303042_1	91.9971671389	2	-	CAPS	*Hin4*II	812	665,147	Not mapped
136	Contig_303042_2	91.9971671389	1	-	CAPS	*BseR*I	798	418,38	Not mapped
142	Contig_1561286_1	92.0679886686	10	1	CAPS	*Bve*I	1190	850,343	Not mapped
143	Contig_1563733_1	92.0679886686	1	-	CAPS		Not suitable for CAPS assay		Not mapped

[1]- Order refers to the marker order in Appendix Table 3.

[2]- Genetic position refers to the genetic positions on the physical map of barley IBSC (2012).

[3]-Number of SNPs detected within the amplicon.

[4]- Number of deletions detected within the amplicon.

[5]- Selected enzyme used for CAPS assay.

[6]- Expected fragment size after digestion.

[7]- Mapped/Not mapped means the marker was used/not used for saturation the target interval.

3.1.3 High-resolution genetic mapping of the gene *albostrians*

During the attempt of marker saturation of the *albostrians* locus the genetic size of the target interval could be reduced to 6.05 cM between markers Contig_220966 and 1_0169, flanking an estimated physical distance of around 10.56 Mbp based on the information provided by the physical map of barley (International Barley Genome Sequencing Consortium, 2012). In total, at least 51 putative genes were predicted to be present in this physical region. Not surprisingly, the obtained recombinations based on mapping in 91 F_2 genotypes did not provide sufficient genetic resolution for immediate delimitation of the gene by map-based cloning. Nevertheless, the *albostrians* gene resided in a region of relatively high recombination frequency (physical to genetic distance ratio= ~ 1.75 Mb/cM around the *albostrians* locus while the average value is 2 Mb/cM in telomeric region of the barley genome) (International Barley Genome Sequencing Consortium, 2012; Ariyadasa et al., 2014), thus, map-based cloning of the gene after high-resolution mapping was promising.

Additional 960 F_2 genotypes were grown under greenhouse condition to increase the genetic resolution. All plants were screened using the flanking markers CAPS_2536 and CAPS_2560 (Figure 3-2A) for the identification of recombination events in the target interval. As a result, 142 recombinants were identified and subsequently tested by all markers that were co-segregating with gene *albostrians* in the initial 91 F_2 population. The target interval could be narrowed down from originally 6.05 cM to 1.72 cM between markers Zip_2661 and CAPS_2560, respectively. Meanwhile, development of new markers using Morex WGS sequences was continued for the target interval and the gene *albostrians* could be delimited in a 0.15 cM genetic interval between Zip_2661 and a newly developed CAPS marker Contig_40728_2. Furthermore, a cluster of nine markers was co-segregating with the *albostrians* gene (Figure 3-2B). According to the physical map of barley (International Barley Genome Sequencing Consortium, 2012), the physical distance between the two flanking markers Zip_2661 (physical position: 536.13 Mbp) and Contig_40728_2 (physical position: 536.34 Mbp) was estimated to be at least 0.21 Mbp.

In order to further increase the genetic resolution an additional 1,920 F_2 plants were grown, however, in order to save time, recombination events were searched only in the homozygous mutant (*as/as*) plants of the population. Plants with variegated or

albino seedling phenotype, per definition, are homozygous for the mutant *albostrians* allele, thus the genetic position of the gene can be determined after marker mapping without a subsequent need of F_3 phenotyping. In total, 384 variegated seedlings were selected out of 1,920 F_2 plants and screened by flanking markers Zip_2661 and Zip_2680. This identified a single recombinant MM4205-230_9_51 that was further genotyped by the markers of the target interval. As a result, one marker Zip_2662 co-segregated with the phenotype and separated from the cluster of co-segregating markers by a single recombination event. Thus the gene *albostrians* could be delimited to a 0.06 cM genetic interval between the proximal marker Zip_2661 and a distal cluster of nine co-segregating markers including marker Contig_92279 (Figure 3-2C).

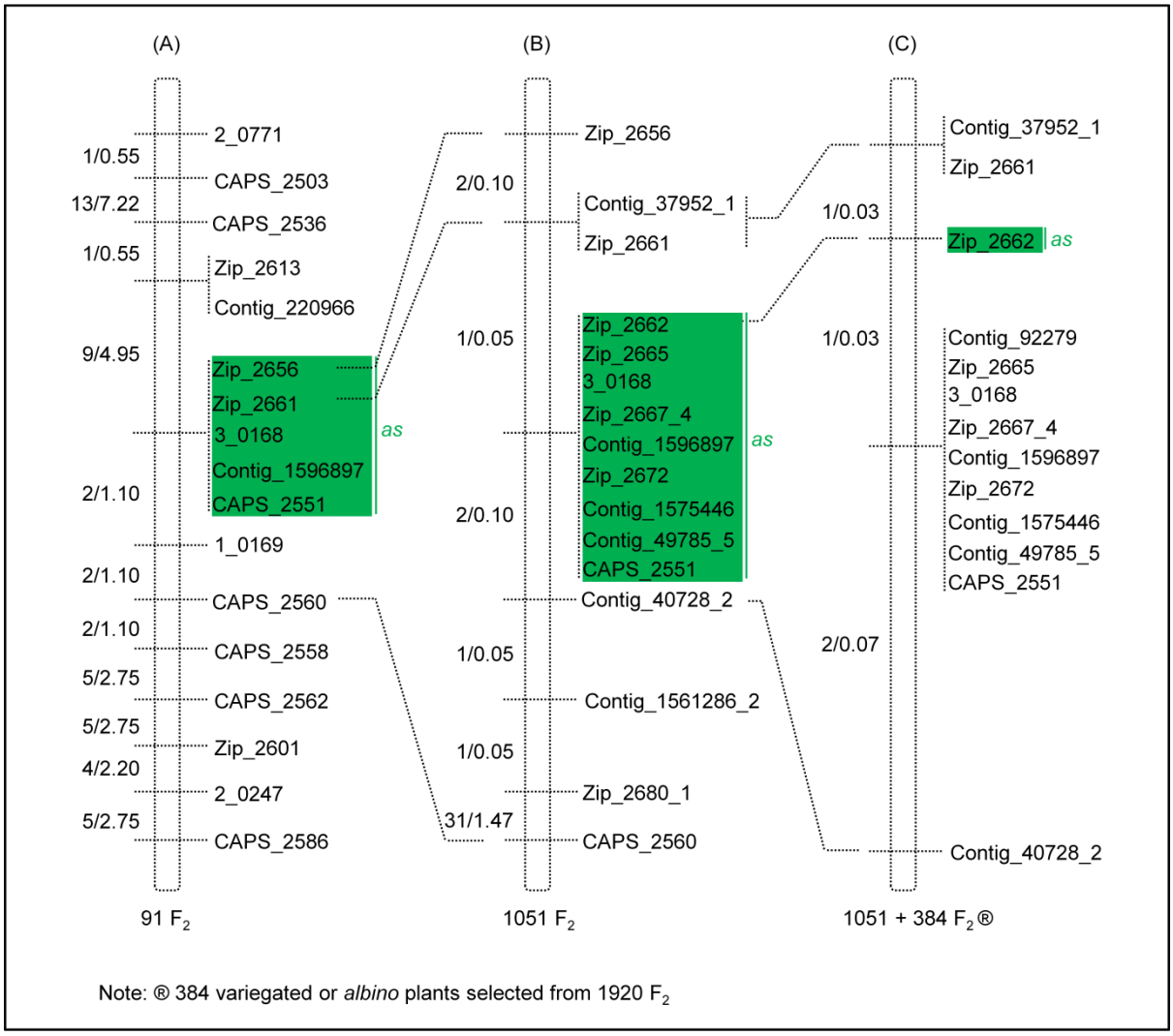

Figure 3-2: High-resolution mapping of the *albostrians* gene. Recombination events and genetic distance (recombinants/cM) between the neighboring markers are shown on the left panel of each genetic map. Markers co-segregating with the gene *albostrians* are highlighted in green background.

(A) Marker saturation around the gene *albostrians* locus. Two new flanking markers, Contig_220966 (or marker Zip_2613) and 1_0169, as well as a cluster of co-segregating markers are identified at the resolution of 91 F_2 genotypes. (B) High-resolution mapping of the gene *albostrains* with 1051 F_2 genotypes. The gene *albostrians* is allocated to a 0.15 cM genetic interval between two flanking markers Zip_2661 (or marker Contig_37952_1) and Contig_40728_2. (C) Further fine mapping the *albostrians* gene using additional 384 variegated or *albino* mutants, which are selected out of 1920 F_2 plants. The gene *albostrians* is allocated to a 0.06 cM genetic interval and in between one marker Zip_2662 co-segregated with the phenotype.

The newly generated genomic resources of barley provided a versatile platform to identify a candidate gene within the identified target interval. On the basis of the barley 'Genome Zipper' information (Mayer et al., 2011) a single gene, carrying the NCBI GenBank accession ID AK366098, was predicted to reside between the flanking markers. Based on the average ratio of genetic to physical distance (0.5 cM/Mbp) predicted by the physical/genetic map for telomeric regions of barley chromosomes (International Barley Genome Sequencing Consortium, 2012; Ariyadasa et al., 2014) the 0.06 cM genetic interval of the *albostrians* locus might comprise an estimated physical distance of less than 120 Kbp. (International Barley Genome Sequencing Consortium, 2012; Ariyadasa et al., 2014).

3.2 Physical mapping of the *albostrians* locus

The *albostrians* gene was delimited to a 0.06 cM genetic interval by high-resolution genetic mapping. As half of the physical map of barley was contained in contigs larger than 904 Kbp (N50 contig length = 904 Kbp) (International Barley Genome Sequencing Consortium, 2012; Ariyadasa et al., 2014) it should be feasible to find the markers flanking the *albostrians* locus present on a single physical BAC contig. Therefore, the flanking markers Zip_2661 and CAPS_2551 as well as their respective co-segregating markers were used to identify a physical map contig that would theoretically contain the gene *albostrians*.

3.2.1 Linking *albostrians* flanking markers to the physical map of barley by sequence comparison and PCR-based screening

Two approaches were used to anchor the closest *albostrians* flanking markers to the physical map of barley (International Barley Genome Sequencing Consortium, 2012). This was either based on *in silico* analysis through sequence comparison (BLAST = Basic Local Alignment Search Tool) or by experimental PCR-screening of a barley

BAC library HVVMRXALLeA derived from the North American six-rowed malting variety 'Morex' (Schulte et al., 2011). All except one marker could be anchored *in silico* (Table 3-4). Marker Contig_1596897 was anchored experimentally by BAC library screening. Two markers Zip_2672 and Contig_1575446 were anchored both by sequence comparison and PCR-based screening.

The two proximally flanking markers Contig_37952_1 and Zip_2661 as well as the co-segregating marker Zip_2662 were allocated to FingerPrinted BAC contig 7112 (FPcontig_7112). FPcontig_7615 was identified by the distally flanking marker 3_0168. All the remaining co-segregating, but the gene *albostrians* distally flanking markers (Zip_2665, Zip_2667_4, Contig_1596897, Zip_2672, Contig_1575446, Contig_49785_5, and CAPS_2551) were allocated to FPcontig_7506 (Table 3-4). Thus instead of a single large BAC contig the identified genetic interval was represented by three individual physical BAC contigs. Based on information provided by the genetically anchored physical map (Mascher et al., 2013a) two additional FP contigs (FPcontig_4483 and FPcontig_44845) were predicted to be located between contigs carrying the flanking markers. In order to identify a possible candidate gene on any of these five BAC contigs the minimally overlapping and non-redundant path of BAC clones was selected for shotgun sequencing (Appendix Table 5).

Table 3-4: Allocating *albostrians* flanking markers to the physical map of barley.

Marker ID	Morex_WGS_Contig	Corresponding BAC	FPcontig	Anchoring Method
Contig_37952_1	Contig_37952	HVVMRXALLeA0035F11	7112	*In silico*
Zip_2661	Contig_65209	HVVMRXALLhB0102L04	7112	*In silico*
Zip_2662	Contig_104939	HVVMRX83KhA0187H24	7112	*In silico*
Zip_2665	Contig_263534	HVVMRXALLmA0412B01	7506	*In silico*
3_0168	Contig_137310	HVVMRXALLrA0254F14	7615	*In silico*
Zip_2667_4	Contig_65509	HVVMRXALLmA0412B01	7506	*In silico*
Contig_1596897	Contig_1596897	HVVMRXALLeA0275D08	7506	BAC screening
Zip_2672	Contig_1575446	HVVMRXALLeA0135I17	7506	*In silico*/BAC screening
Contig_1575446	Contig_1575446	HVVMRXALLeA0135I17	7506	*In silico*/BAC screening
Contig_49785_5	Contig_49785	HVVMRXALLmA0055G09	7506	*In silico*
CAPS_2551	Contig_49785	HVVMRXALLMA0055G09	7506	*In silico*

3.2.2 Sequence analysis of BAC contigs identified by markers that are flanking or co-segregating with the *albostrians* locus

A total number of 60 minimal tiling path (MTP) BAC clones was selected from 5 FPcontigs (Appendix Table 5) for high-throughput short read shotgun sequencing. The obtained sequencing reads were assembled with CLC assembly cell 3.2.2 (http://www.clcbio.com/) after quality trimming and vector / adaptor removal. All-against-all BLAST searches of individual BAC assemblies revealed that FPcontig_7112 and FPcontig_7615 were overlapping by BAC clones HVVMRXALLrA0395M21 and HVVMRXALLmA0230A06. No overlaps could be identified for the remaining three physical contigs (FPcontig_4483, FPcontig_7506 and FPcontig_44845). The proximal flanking marker Zip_2661 and co-segregating marker Zip_2662 were located on the same BAC clone HVVMRXALLrA0395M21 (FP Contig_7112) (Figure 3-3B). The distally flanking marker 3_0168 was located on a contiguous sequence scaffold of FPcontig_7615 thus markers flanking the locus *albostrians* on both sides were allocated to a single sequence scaffold represented by the MTP sequence of the physical contigs FPcontig_7112 and FPcontig_7615. A new marker Contig_92279 could be developed based on the sequence information of BAC HVVMRXALLrA0395M21, which contained the co-segregating marker Zip_2662 as well as the proximal flanking marker Zip_2661. Mapping of marker Contig_92279 revealed its distal allocation in regard of the gene *albostrians*, thus a physical region of about 46 Kbp was characterized genetically and physically to comprise the *albostrians* locus (Figure 3-3C). The ratio of physical to genetic distance of this region equalled about 0.77 Mb / cM, indicating that the *albostrians* gene was located in a 'hot spot' of recombination (Kunzel et al., 2000). In order to identify (a) possible candidate gene(s), gene models were predicted for non-repetitive sequences of the target interval through sequence comparison to annotated gene models defined on the Morex reference WGS assembly (International Barley Genome Sequencing Consortium, 2012). Only one single gene *MLOC_670* (GenBank accession ID AK366098) could be identified within the target interval (Figure 3-3C). Re-sequencing of the gene *MLOC_670* revealed a 4 bp deletion in mutant M4205 compared to wild type. This mutation was predicted to induce a frame-shift in the reading frame and as a consequence a premature stop codon in the second exon of the gene (Figure 3-3D) which ultimately might result in loss of function of the gene in the mutant. Therefore, the identified gene was a strong candidate for representing the *albostrians* gene.

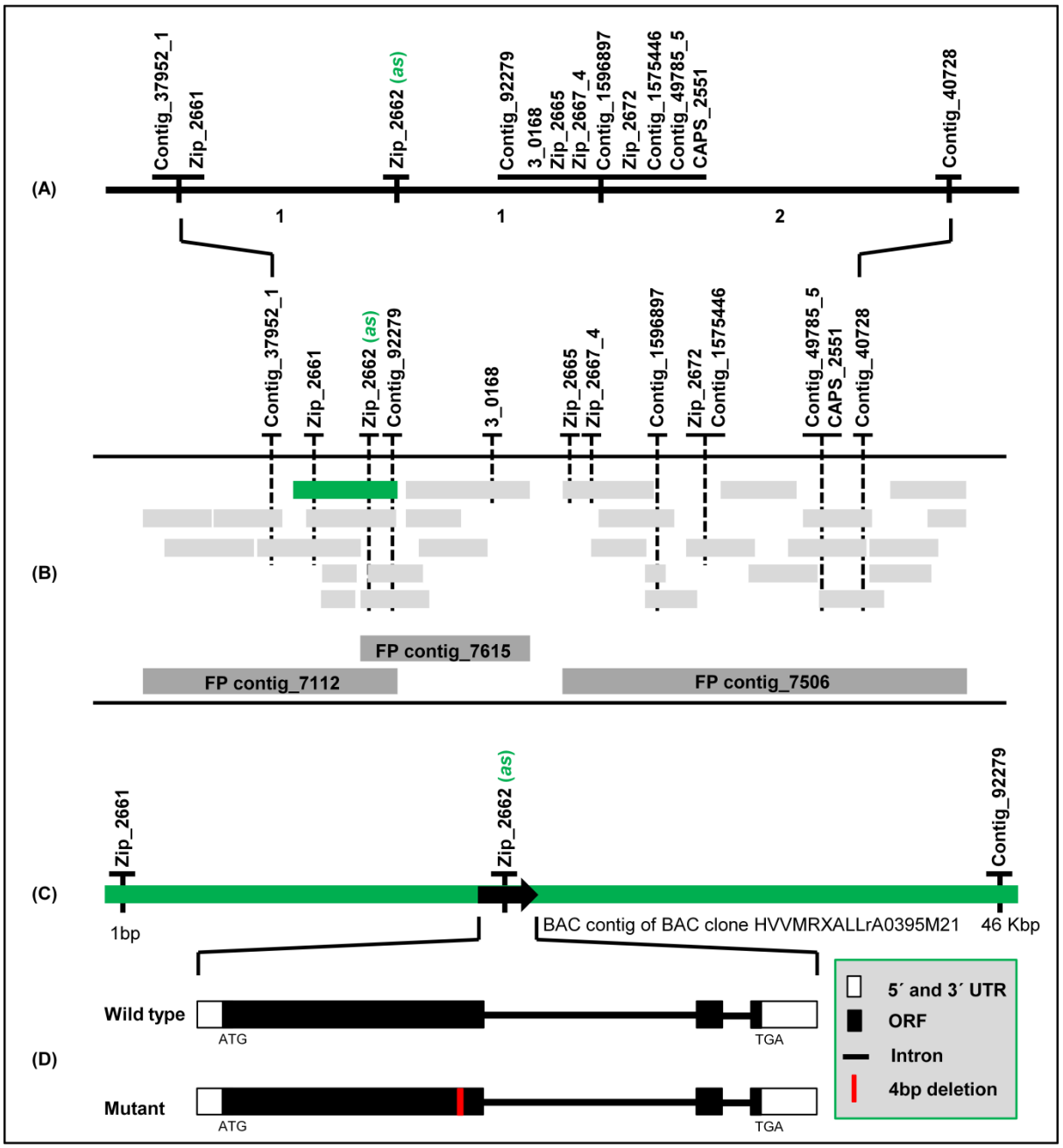

Figure 3-3: Summary of genetic and physical mapping of the *albostrians* gene. (A) High-resolution genetic mapping of the *albostrians* gene. One co-segregating marker (Zip_2662) is identified and the recombination between contiguous markers is indicated by the numbers below. (B) Physical anchoring of markers to the sequenced MTP BACs. The two flanking markers Zip_2661 and Contig_92279 as well as the co-segregating marker Zip_2662 are located on the same BAC clone HVVMRXALLrA0395M21 (FPcontig_7112) as indicated in green color. (C) Gene prediction based on repeat masked BAC assemblies of the target interval through alignment to barley gene models published by IBSC (2012). A single gene, *MLOC_670* (GenBank accession ID: AK366098) as indicated by the arrow, is identified within the target interval between two flanking markers Zip_2661 and Contig_92279, flanking a physical distance of around 46 Kbp. (D) Structure of the gene *MLOC_670*: The mutant allele harbors a 4 bp deletion (red bar) near the end of the first exon as compared to the wild type allele found in genotype Morex.

3.3 Structural and functional annotation of the *albostrians* candidate gene

The *albostrians* candidate gene, *MLOC_670*, was identified through map-based cloning. It contained three protein coding exons as revealed by RT-PCR, followed by sequence alignment to genomic DNA (Figure 3-4A) with a total length of 1377 bp leading to a putative protein length of 459 amino acids (AA) upon translation. The three exons had a length of 1152 bp, 171 bp, and 54 bp, respectively. Two introns with a length of 1140 bp and 162 bp resided between exon1 and 2 and exon 2 and 3, respectively. Orthologous genes of *MLOC_670* in rice and *Brachypodium* showed a similar gene structure to the barley gene (Figure 3-4B&C).

3.3.1 Functional annotation of the *albostrians* candidate gene

The *MLOC_670* gene was annotated as 'CCT motif family protein' (Cockram et al., 2012; International Barley Genome Sequencing Consortium, 2012). The CCT domain was originally described as a 43 AA region of homology to the *Arabidopsis thaliana* proteins CONSTANS (CO), CO-LIKE, and TIMING OF CAB1 (TOC1) (Putterill et al., 1995; Strayer et al., 2000). Gene ontology analysis (Jones et al., 2014) revealed that the conserved CCT domain of *MLOC_670* ranged from AA residue position 405 to 447 (predicted protein sequence started from the methionine encoded by the start codon). The gene belongs to the fourth group of the *CCT motif family* (*CMF*) genes, which possess an intron after residue 37 of the conserved domain as found in other members of the gene family from *Poaceae* (*CMF3*, *CMF7*, *CMF9*) and Arabidopsis (*AtCMF3*, *AtCMF9*, *AtCMF11*, *AtCMF14*) (Cockram et al., 2012). The CCT motif was predicted to be a nuclear localization signal (Robert et al., 1998), and all of the investigated CCT motif-containing proteins were shown to be transcription factors (Sun et al., 2001). In addition*, in silico* analysis by a list of online prediction tools showed that the putative ALBOSTRIANS protein contained an N-terminal chloroplast transit peptide and thus is potentially localized to the chloroplast (Table 3-5).

A BLAST sequence search of the MLOC_670 protein sequence against the Arabidopsis genome identified its putatively orthologous gene *chloroplast import apparatus 2* (*cia2*). Arabidopsis *cia2* mutant plants exhibited a pale green phenotype (Sun et al., 2001). Functional studies on the *CIA2* gene showed that it codes for a putative transcription factor that was mainly binding promoter regions of genes encoding chloroplast protein translocon components and chloroplast ribosomal

proteins (Sun et al., 2009). Therefore, the protein encoded by *MLOC_670* is postulated to represent as well a transcription factor with potential regulating activities on the genes encoding translocon proteins and/or chloroplast ribosomal proteins in barley.

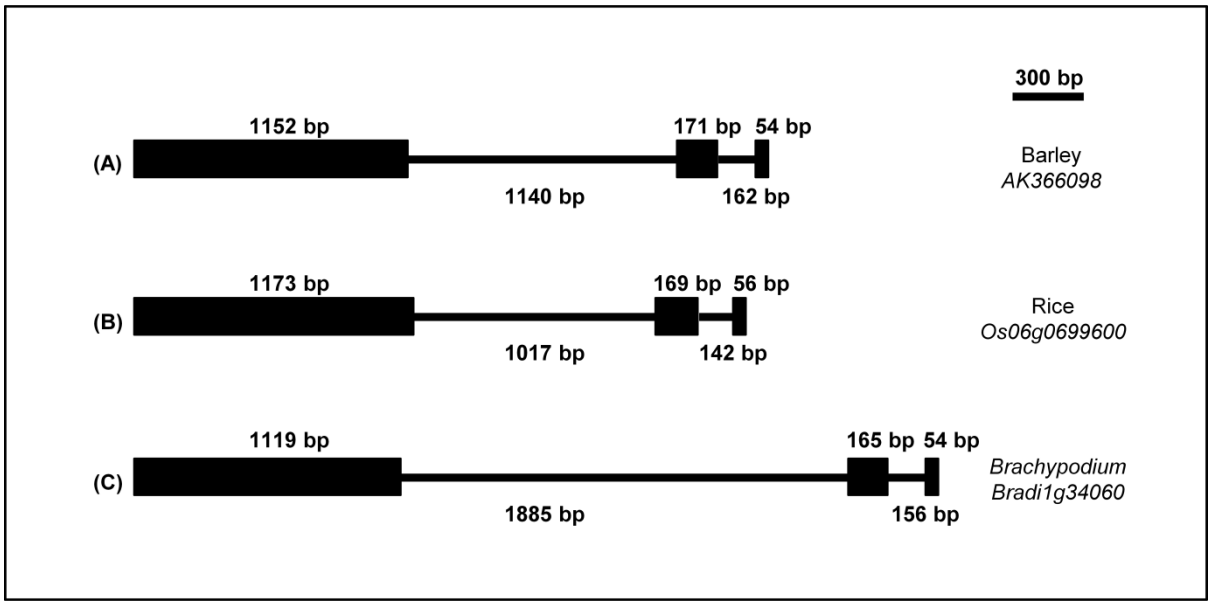

Figure 3-4: Structural annotation of *MLOC_670* and its orthologs in rice and *Brachypodium*. (A) The *albostrians* candidate gene, *MLOC_670*, contains three exons (indicated by rectangles). Gene structure of the orthologs in rice and *Brachypodium* are presented in (B) and (C), respectively. The length of the exons/introns is indicated by the numbers above/below each of the schematic drawings, respectively.

Table 3-5: *In silico* prediction of subcellular localization of the ALBOSTRIANS protein.

Online Tools	cTP Score[A]	mTP Score[B]	SP Score[C]	Others	cTP	cTP-length	Predicted Location
TargetP	0.742	0.023	0.009	0.339	Yes	82	Plastid
PredSL	0.749	0.002	0.004	-	Yes	83	Plastid
ChloroP	0.525	-	-	-	Yes	82	Plastid
Predotar	0.33	0.01	-	0.64	-	-	Possibly plastid
WoLFPSORT	8	-	-	5	-	-	Possibly Plastid

[A]- cTP represents chloroplast transit peptides.

[B]- mTP represents mitochondrial transit peptides.

[C]- SP represents secretory pathway.

3.3.2 *In silico* expression analysis of the *albostrians* candidate gene

To further understand the function of the *albostrians* candidate gene, the expression profile of *MLOC_670* was examined on the basis of publicly available gene expression data obtained for eight tissues or developmental stages (Figure 3-5) (International Barley Genome Sequencing Consortium, 2012) (http://apex.ipk-

gatersleben.de/apex/f?p=284:20:0::NO::P20_GENE_NAME:AK366098). Expression of the *MLOC_670* gene was detected across all the eight tissues/stages. The highest level of expression was found in young barley tillers. In contrast, the embryo exhibited the lowest expression level.

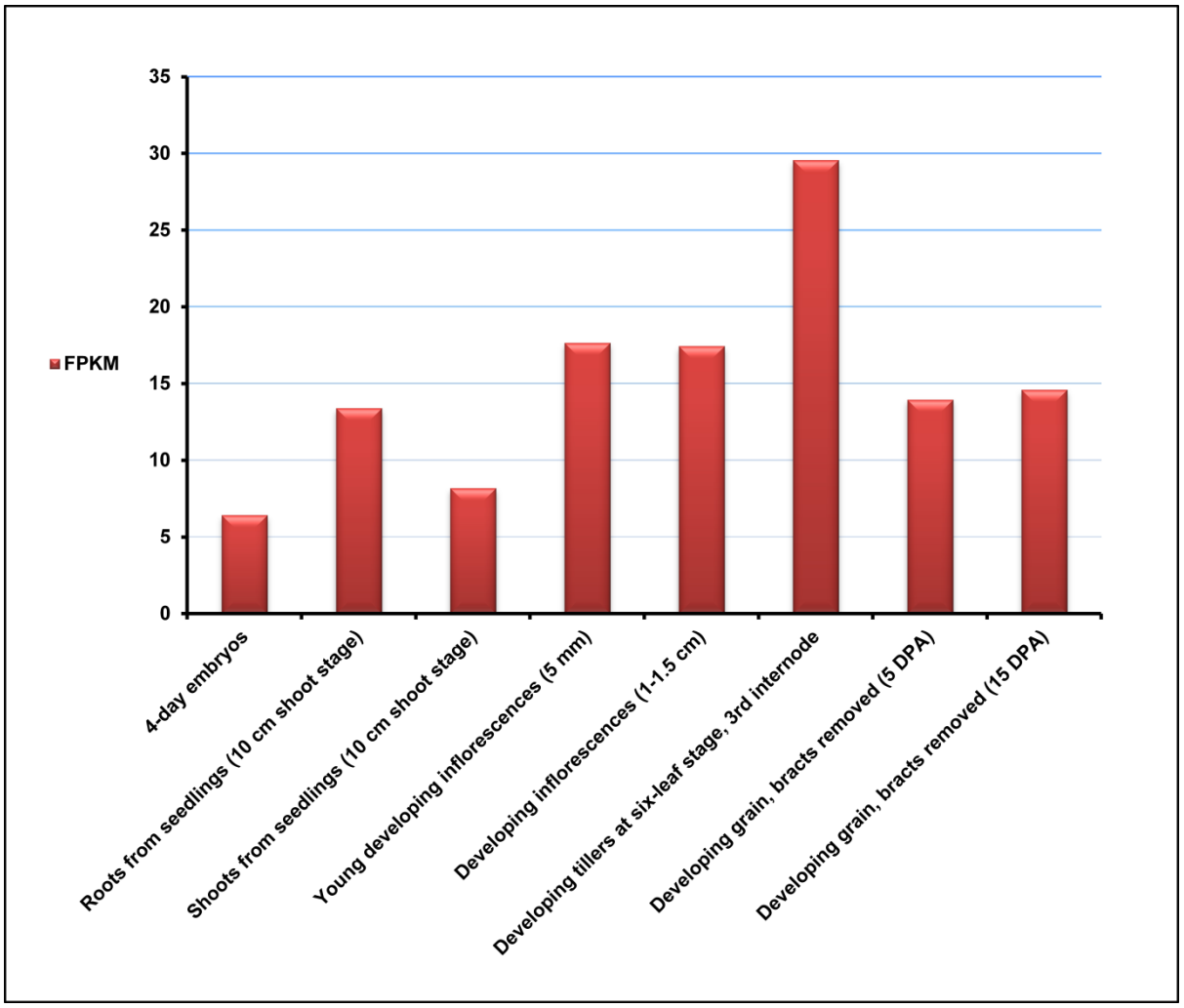

Figure 3-5: Expression of the *MLOC_670* gene in barley. The expression level is given as fragments per kilobase of exon per million reads mapped (FPKM) across eight different tissues or developmental stages. Gene expression could be detected in all tissues with highest gene expression level in young plants shifting from vegetative to generative growth stages. The data is taken from the International Barley Genome Sequencing Consortium (2012).

3.4 Functional validation of the *albostrians* candidate gene

The gene *MLOC_670* was identified as the only candidate gene responsible for the variegated phenotype of the *albostrians* mutant. Compared to WT plants, the mutant carried a 4 bp deletion leading to a premature stop codon in the gene. Hence, the truncated protein in the mutant might have lost its function completely or, might

exhibit residual biological activity. Four independent approaches were implemented to validate the candidate gene. First, a TILLING population was screened for identification of independent novel mutated alleles. Second, a series of independent mutants segregating for an *albino* phenotype were screened for mutations in the candidate gene. Third, a transgenic approach to induce independent knock-out mutants by a TAL Effector Nuclease (TALEN) construct targeting the *albostrians* candidate gene was initiated. Fourth, another transgenic approach was followed to recover the mutant phenotype by functional complementation of a mutant genotype with a WT copy of the *albostrians* candidate gene. The latter two transgenic experiments could not be finalized in duration of the project and results are still pending. The results of the initial two strategies are shown in the following.

3.4.1 Functional validation by TILLING

TILLING is a reverse genetics approach as developed by Colbert et al. (2001). This strategy utilizes traditional mutagenesis followed by high throughput mutation discovery (McCallum et al., 2000). It has been proven to be a powerful tool for functional analysis of barley genes (Gottwald et al., 2009; Gawronski et al., 2014; Mascher et al., 2014; Yang et al., 2014).

Two primer pairs were designed to cover all three exons of the *albostrians* candidate gene (Table 3-6) and used to screen a total of 7,979 M_2 plants. This revealed 42 EMS-induced mutations, including 20 synonymous and 20 non-synonymous mutations, one 9 bp deletion and one mutation leading to a premature stop codon (Table 3-7). M_3 families of the identified M_2 plants were phenotyped and all individuals were genotyped for the presence of the expected mutation by re-sequencing the *albostrians* candidate gene. Six homozygous mutant M_3 plants could be identified in M_2-TILLING family 6460-1 segregating for the premature stop codon mutation. All of these six plants showed a complete *albino* phenotype (Figure 3-6A) supporting the hypothesis that the identified candidate gene is indeed involved in chloroplast development. Five heterozygous plants and two homozygous wild type plants of this M_2-TILLING family 6460-1 were propagated to produce M_4 seeds. Two hundred forty-five M_4 plants derived from the five heterozygous M_3 plants were phenotyped and genotyped at the *albostrians* candidate gene. Consistently, all homozygous mutant genotypes grew into purely white (*albino*) seedlings, with a

single exception, where a narrow green stripe was observed on the first leaf of the M_4 seedling, indicating a link to the variegated phenotype originating from the original *albostrians* mutation. The second leaf of this plant, however, turned again into totally *albino* phenotype (Figure 3-6B). M_4 plants with heterozygous (117 plants) or homozygous wild type (65 plants) genotype at the *albostrians* candidate gene (Table 3-8) were always completely green. The results were tested for statistical significance by a chi-squared test ($0.7< p <0.8$) indicating that the *albino* phenotype of this mutant family was controlled by a single recessive gene, hence the phenotypic effect was completely linked to the mutation in the *albostrians* candidate gene and thus was not due to background mutations induced by EMS, which occasionally can lead to segregation of *albino* phenotype in independent M_3 (or higher generation) TILLING families. This was further supported by the fact that all 34 plants obtained from the two wild type M_3 families all showed a fully green phenotype which was consistent with the genotype at the *albostrians* locus. This data strongly support the view that the identified candidate *albostrians* gene *MLOC_670* is the functional gene underlying the mutant phenotype of the original *albostrians* mutant genotype M4205. Thus the gene was named *HvAs* (for: *Hordeum vulgare Albostrians*). Alternatively, in reference to its putative ortholog of Arabidopsis, where a gene function was reported for the first time, the name *HvCIA2-like* may be used.

The original *albostrians* mutant carried a 4 bp deletion at nucleotide position 1123 - 1126 (count from adenine of start codon as +1). As a consequence, a truncated (shorter) protein with only 392 AA would be the result in case of translation. The identified TILLING mutant (M_2-TILLING family 6460-1) carried a premature stop codon at nucleotide position 928 resulting, in case of translation, in an even shorter protein of a length of only 309 AA. Thus a putative protein of the *HvAs* allele of TILLING mutant is putatively 83 AA shorter than the protein of the original mutant M4205 (Figure 3-7). This size difference may help to explain the observed phenotypic differences of the two independent mutants. Whereas seedlings of the *albostrians* mutant may either be green, variegated (alternating longitudinal green and white sectors) or completely *albino*, homozygous mutant progeny of the TILLING mutant were predominantly *albino* or in very rare cases may exhibit small sectors of green in otherwise *albino* plantlets. This observation may be explained by the size differences of the two truncated proteins: The longer protein resulting from the M4205 *HvAs*

allele would still, at a certain threshold, allow normal chloroplast development; the even shorter putative protein of the TILLING mutant would have lost this residual activity, thus normal chloroplast development in homozygous mutant genotypes is impaired.

Table 3-6: Primers used for screening the TILLING population.

Primer ID	Length (bp)	GC%	T_m (°C)	Sequence	Region	Product Size (bp)
AK366098_Exon1_F5	20	50.00	59.15	GGGTCCAGATTGATTCATCC	Exon 1	1376
AK366098_Exon1_R5	20	50.00	61.74	GCAGTGCAGGCATTTCAATC	Exon 1	
AK366098_Exon3_F	20	50.00	59.96	ATCAGGGAGCATGGTTTACG	Exon 2&3	596
AK366098_Exon3_R	20	50.00	60.02	AGCCGTCATCTGCTTCACTT	Exon 2&3	

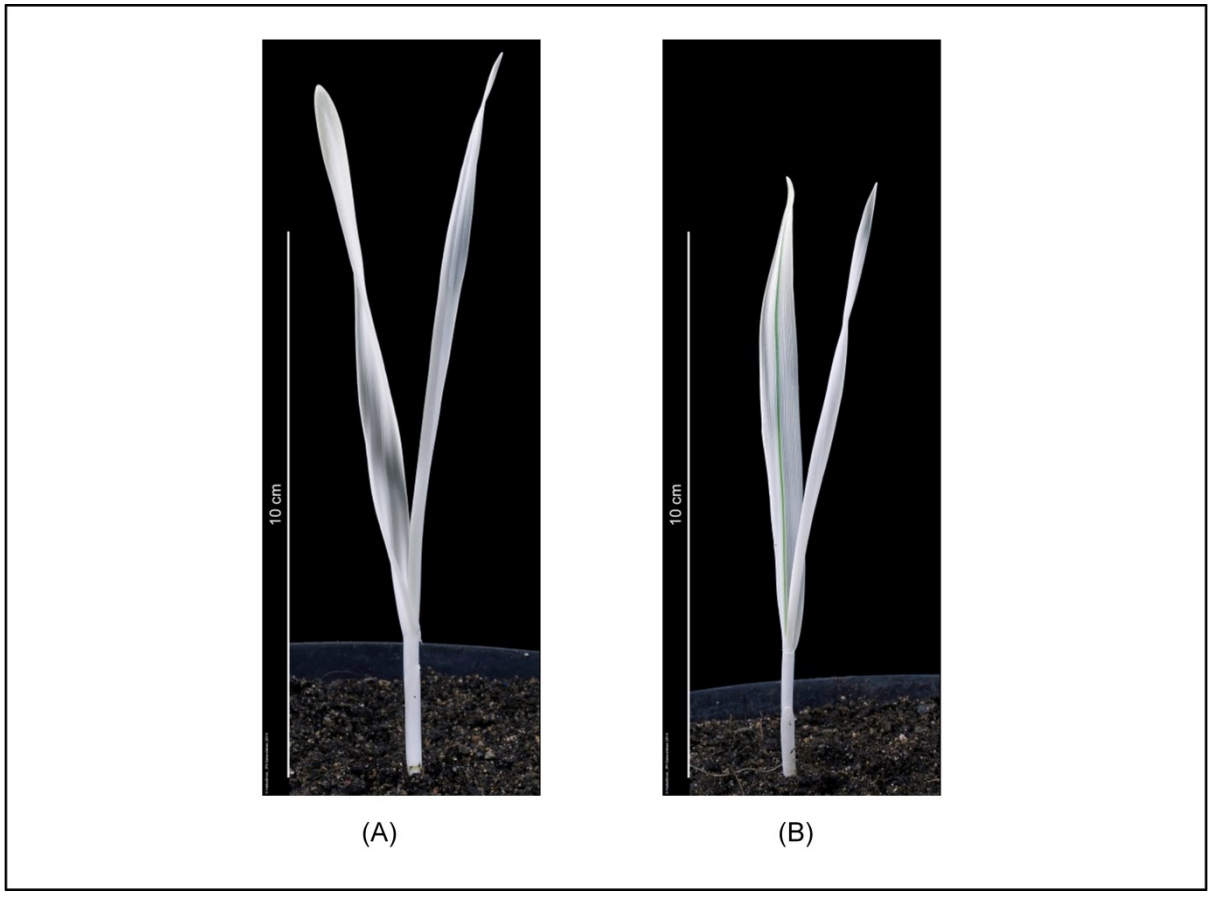

Figure 3-6: Phenotypic analysis of homozygous mutant M_3 and M_4 plants of M_2-TILLING family 6460-1 carrying a premature stop mutation in the gene *HvAs*. (A) Typical seedling phenotype of M_3 and M_4 plants homozygous mutant for the gene *HvAs*. (B) A single M_4 mutant of M_3-TILLING family 6460-1_9 exhibits a slightly variegated first leaf showing a small green sector which may indicate low residual functionality of the mutant protein.

Table 3-7: Summary of the identified TILLING mutants.

Plant Family ID[1]	SNP Position[2]	SNP	Original[3]	M_2/M_4 Status[4]	AA Change	Note[5]
10580-1	30	C/T	AT**C**	Heterozygote	Ile/Ile	Exon 1
3283-1	60	G/C	GC**G**	Heterozygote	Ala/Ala	Exon 1
12436-1	72	C/T	GC**C**	Heterozygote	Ala/Ala	Exon 1
9688-1	114	G/A	CC**G**	Homozygote	Pro/Pro	Exon 1
3227-1	130-138	Deletion	**TCCTCGGCG**	Homozygote	Ser Ser Ala/---	Exon 1
7924-1	156	C/T	AA**C**	Homozygote	Asn/Asn	Exon 1
9414-1	168	C/T	GC**C**	Heterozygote	Ala/Ala	Exon 1
14946-1	222	C/T	AC**C**	Heterozygote	Thr/Thr	Exon 1
9878-1	279	G/T	GG**G**	Homozygote	Gly/Gly	Exon 1
8265-1	286	G/A	**G**CG	Homozygote	Ala/Thr	Exon 1
6646-1_24_1	286	C/T	GA**C**	Homozygote	Asp/Asp	Exon 1
9825-1	286	C/T	**C**TC	Homozygote	Leu/Phe	Exon 1
13029-1	367	C/T	**C**CC	Heterozygote	Pro/Ser	Exon 1
10459-1	367	C/T	CT**C**	Homozygote	Leu/Leu	Exon 1
3779-1	427	C/T	**C**CG	Homozygote	Pro/Ser	Exon 1
12772-1	435	C/T	AG**C**	Homozygote	Ser/Ser	Exon 1
13269-1	504	C/T	TC**C**	Heterozygote	Ser/Ser	Exon 1
10799-1	514	C/T	**C**CG	Homozygote	Pro/Ser	Exon 1
9920-1	634	G/A	**G**CC	Homozygote	Ala/Thr	Exon 1
13239-1	641	G/A	G**G**G	Heterozygote	Gly/Glu	Exon 1
2889-1	642	G/A	GG**G**	Heterozygote	Gly/Gly	Exon 1
2920-1	659	G/A	G**G**C	Heterozygote	Gly/Asp	Exon 1
7669-1	664	C/T	**C**TC	Homozygote	Leu/Phe	Exon 1
3698-1	667	A/T	**A**GT	Homozygote	Ser/Cys	Exon 1
6912-1	695	C/T	A**C**T	Heterozygote	Thr/Ile	Exon 1
12113-1	730	C/T	**C**CC	Heterozygote	Pro/Ser	Exon 1
8222-1	774	C/T	CA**C**	Heterozygote	His/His	Exon 1
12189-1	821	G/A	A**G**C	Heterozygote	Ser/Asn	Exon 1
4394-1	847	C/T	**C**CA	Heterozygote	Pro/Ser	Exon 1
12710-1	886	G/A	**G**CG	Homozygote	Ala/Thr	Exon 1
10022-1	909	G/A	GA**G**	Homozygote	Glu/Glu	Exon 1
6460-1	928	A/T	**A**AG	Heterozygote	Lys/stop codon	Exon 1
10996-1	1009	G/A	**G**AT	Heterozygote	Asp/Asn	Exon 1
13603-1	1009	G/A	**G**AT	Heterozygote	Asp/Asn	Exon 1
10972-1	1032	G/A	AA**G**	Heterozygote	Lys/Lys	Exon 1
3812-1	1062	C/T	CT**C**	Heterozygote	Leu/Leu	Exon 1
11403-1	2358	G/A	A**G**C	Homozygote	Ser/Asn	Exon 2
13059-1	2379	G/A	AA**G**	Homozygote	Lys/Lys	Exon 2
14134-1	2385	G/A	AA**G**	Heterozygote	Lys/Lys	Exon 2
3005-1	2430	G/A	CG**G**	Heterozygote	Arg/Arg	Exon 2
4217-1	2434	G/A	**G**TG	Homozygote	Val/Met	Exon 2
11797-1	2441	C/T	G**C**C	Heterozygote	Ala/Val	Exon 2

[1]- Plant family ID refers to M_2 TILLING families with one exception 6646-1_24_1 in regarding to M_4.

[2]- Coordinates based on genomic sequence of cv. Barke. The adenine of start codon is counted as position +1.

[3]- The changed nucleotides were marked in bold.

[4]- Plant family 6646-1_24_1 has no M_3 seeds in stock.

[5]- Defined regions of the mutation.

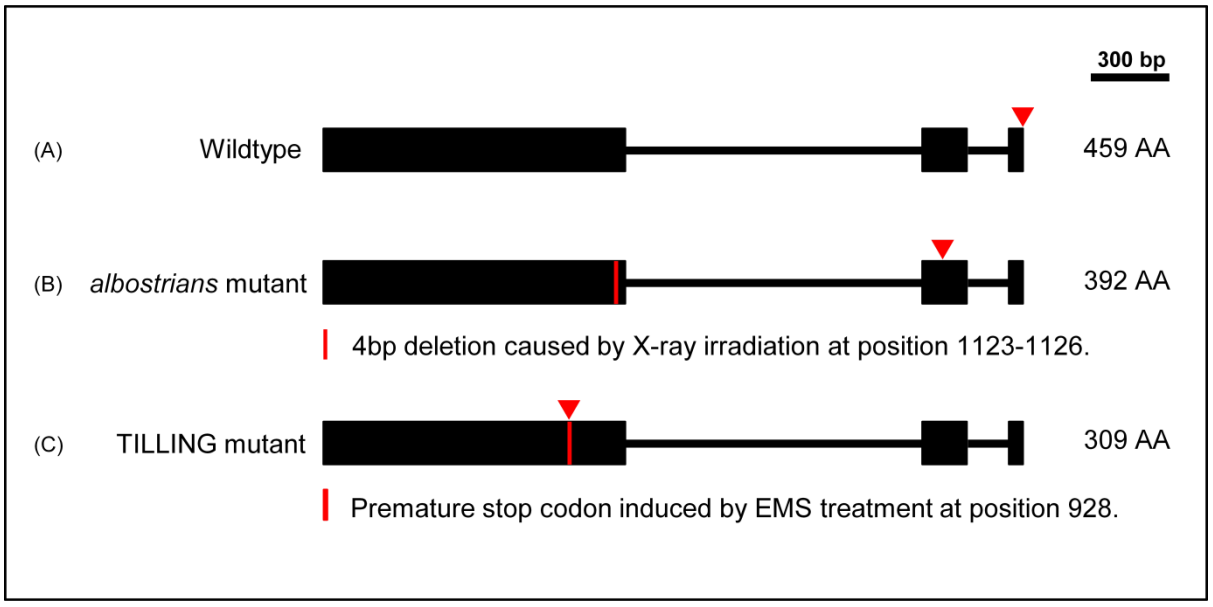

Figure 3-7: Comparison of the structure of WT and MT *HvAs* alleles. (A) Structure of the WT *HvAs* gene, the coding sequence with a total length of 1377 bp. (B) The original *albostrians* mutant M4205 carrys a 4 bp deletion at the tail of the first exon caused by x-ray irradiation. The truncated coding sequence is 1176 bp. (C) The identified M_2-TILLING mutant 6460-1 shows an immature stop codon at nucleotide position 928. The stop codon of the WT/MT *HvAs* alleles is indicated by a red triangle. The length of deduced protein from three versions of the *albostrians* gene is indicated on the right side.

Table 3-8: Analysis of M_4 plants of M_2_TILLING family 6460-1.

M_3 Plant Family ID	Wild Type (M_4)	Heterozygous (M_4)	Mutant (M_4)
6460-1_4	12	30	16
6460-1_9	17	18	18
6460-1_11	8	24	10
6460-1_15	14	13	6
6460-1_18	14	32	13
Total	**65**	**117**	**63**

3.4.2 Mutant analysis indicated *HvAs* is different from other '*albino* genes' in the NordGen mutant collection

Mutant collections proved to represent an invaluable tool in barley functional gene analysis. Prominant examples of how historical mutants supported the identification of candidate genes were provided by the cloning of row-type controlling genes *six-rowed spike 1* (*Vrs1*) and *six-rowed spike 4* (*Vrs4*) (Komatsuda et al., 2007; Koppolu et al., 2013) as well as the stem architecture gene *many-noded dwarf* (*mnd*, Mascher et al., 2014). The original *albostrians* mutant, and especially, subsequently the identified M_2-TILLING mutant 6460-1, both may lead to *albino* seedling phenotype. It might be possible therefore, that historical *albino* mutant collections could reveal

additional independent mutant alleles of the gene *HvAs*. In this context, 28 *albino* mutant plant families were obtained from NordGen and analyzed. Re-sequencing of the gene *HvAs* in *albino* seedlings from each of the 28 plant families did not reveal new mutant alleles of the *HvAs* gene. Furthermore, RT-PCR showed that transcription of the *HvAs* gene in green and *albino* plants from each of the 28 plant families, excluding the possibility of promoter mutations that could negatively impact the expression of the *HvAs* gene (data not shown). As a conclusion, the analysis of 28 *albino* mutant families from a historical mutant collection did not reveal any additional independent mutant allele of the gene *HvAs*.

3.5 Subcellular localization of ALBOSTRIANS protein

Based on the functional annotation of the gene *HvAs* the expressed putative protein was predicted to be transported to the chloroplast and the nucleus (see section 3.3.1). A variety of experimental approaches have been developed for studying sub-cellular localization of proteins in plant cells. This may include either the use of fluorescently labeled fusion proteins or immunolabeling by native protein-specific or tag-specific antibodies. In a first attempt of determining the subcellular localization of the ALBOSTRIANS protein, C-terminal GFP-fusion proteins of wild-type and the two mutated alleles (the original *albostrians* mutant M4205 and M_2-TILLING mutant 6460-1) were generated. The three fusion constructs, together with one plastid marker pt-rk-CD3-999, respectively, were then subjected to transient expression experiments using biolistic bombardment of barley leaf segments with vector-coated gold particles. The plastid marker pt-rk-CD3-999 (Nelson et al., 2007) was detected with artificially assigned orange fluorescence and the chlorophyll fluorescence of the chloroplasts in the mesophyll cells was indicated with red color. Cells expressing the GFP-only (non-fusion GFP control), showed green fluorescence associated with the nucleus and the cytoskeleton / cytoplasm strands but not with proplastids (Figure 3-8A). Transient expression of the wild-type fusion construct was detected as green fluorescence in leaf epidermal cells where the signal was associated with nucleus and plastids (Figure 3-8B). The mutated proteins displayed the same subcellular localization as the wild-type protein (Figure 3-8C&D). In total, 30-50 cells for each of the three fusion constructs were checked for presence of both green and orange fluorescence. Taken together, the wild-type ALBOSTRIANS and its two mutated forms displayed a compartmentalized accumulation in the plastid and might also be

localized to the nucleus. Nevertheless, to investigate if the signal allocated to in the nucleus is a consequence of the GFP fusion to ALBOSTRIANS or of unspecific nuclear targeting of GFP as was reported in other systems before (Seibel et al., 2007), evidence through independent methods such as immno-labeling with an ALBOSTRIANS specific antibody will be needed to confirm the current result.

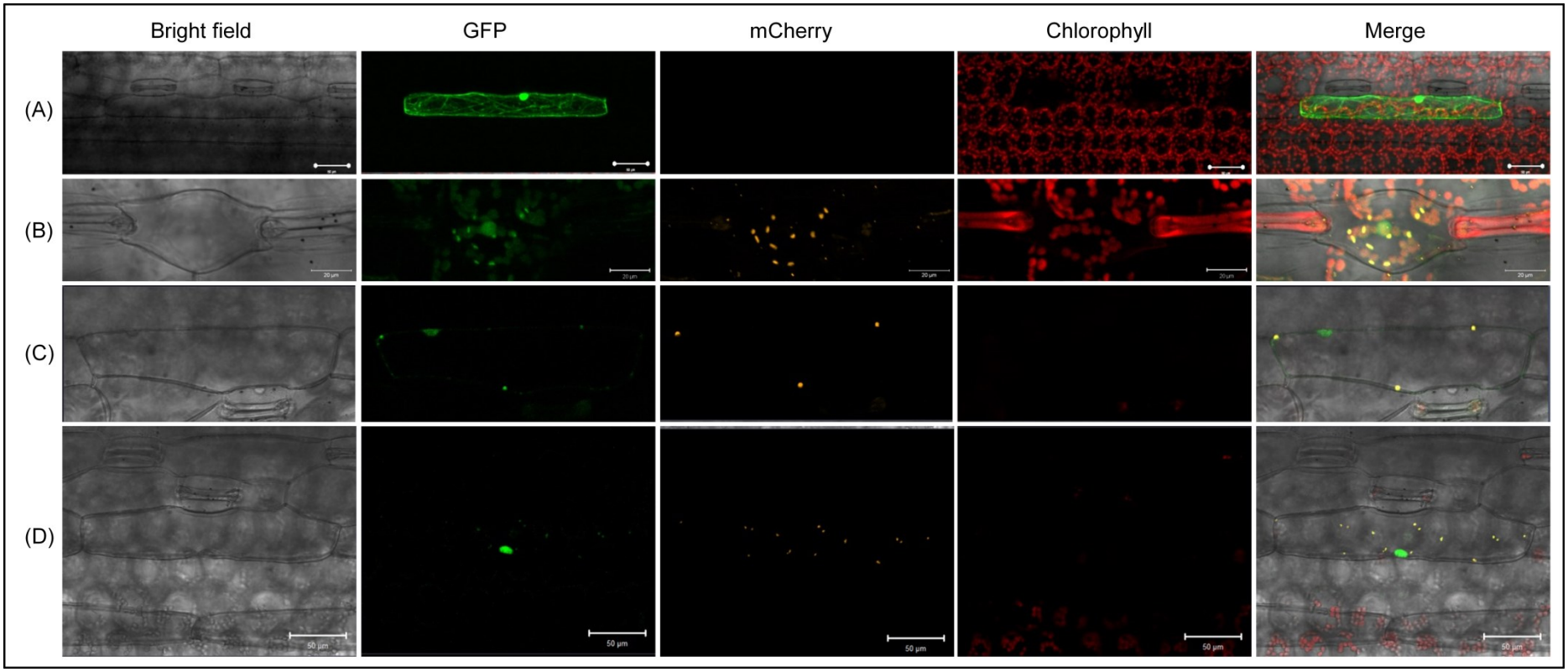

Figure 3-8: Subcellular localization of HvAs-GFP fusion proteins. (A) Particle bombardment of gold-coated pSB179 construct (expressing GFP only) as control. The green fluorescence is distributed in the nucleus and the cytoskeleton / cytoplasm. (B) Particle co-bombardment of both gold-coated plastid marker pt-rk-CD3-999 and wild-type fusion construct pSB179-HvAs_WT. The bright field, GFP, mCherry and chlorophyll signals are displayed as individual channels and the merged channels are shown on the rightmost panel. The GFP fluorescence signals, mCherry fluorescence signals and chlorophyll auto-fluorescence signals are indicated with green, orange and red artificial color, respectively. The green fluorescence of the wild-type fusion protein accumulated in plastids and nucleus of the barley epidermal cells, i.e. the ALBOSTRIANS protein might be dually targeted to both plastid and nucleus. As demonstrated for the wild-type, subcellular localization of the two mutated proteins pSB179-HvAs_M4205 (C) and pSB179-HvAs_TILLING (D) were also investigated by co-bombardment with the plastid marker pt-rk-CD3-999, respectively. The mutated proteins displayed the same subcellular localization as the wild-type protein. The first leaf of 10-day-old barley seedlings was used for particle bombardment. The fluorescence was checked 24 hours after bombardment.

3.6 Phylogenetic analysis of the gene *HvAs* in context of the orthologous and paralogous members of the *CMF* gene family

The gene *HvAs* is a member of the *CCT Motif Family* (*CMF*) genes, featured by the presence of only one conserved CCT domain Cockram et al. (2012). In barley this gene family comprises ten genes (Table 3-9), while in other grass species like rice (*Oryza sativa*) and *Brachypodium distachyon*, 14 and 15 *CMF* members were found, respectively (Cockram et al., 2012). Numerous CCT domain-containing genes are involved in the photoperiod pathway (Putterill et al., 1995; Suarez-Lopez et al., 2001; Valverde, 2011). In cereals, natural variation within CCT domain genes is critical in the control of flowering and adaptation after crop domestication (Cockram et al., 2012).

Table 3-9: Members of CCT motif family in barley (*Hordeum vulgare*).

Name[A]	Accession ID	Chromosome	Rice Homologous	Amino Acid	cTP score[B]	CCT Domain[C]
HvCMF1	*MLOC_59697.1*	3HL	*Os01g61900*	415	0.097	-
HvCMF3	*AK357603*	6HS	*Os02g05470*	490	0.865	436 - 479
HvCMF4	*AK375853*	4H	*Os03g04620*	388	0.110	290 - 334
HvCMF5[D]	*AFX97550.1*	1H	*Os05g38990*	221	-	173 - 216
HvCMF6a	*AK250075.1*	NA	*Os05g51690*	344	0.885	292 - 334
HvCMF6b	*AK355694*	NA	*Os05g51690*	328	0.468	276 - 318
HvCMF7	*AK366098*	7HL	*Os06g48610*	459	0.742	403 - 448
HvCMF10	*MLOC_67881.2*	1H	*Os10g32900*	397	0.061	301 - 345
HvCMF11	*MLOC_75450.2*	1HL	*Os10g41100*	295	0.515	220 - 263
HvCMF13	*MLOC_9998.3*	5HL	*Os12g01080*	198	0.772	88 - 131

[A]- The names are on the basis of Cockram et al. (2012).

[B]- cTP score is predicted according to the online tool TargetP. It reflects the reliability of chloroplast localization.

[C]- The CCT domain is predicted according to InterPro (Jones et al., 2014). '-' means no prediction result.

[D]- Only partial protein sequence is available for *HvCMF5* according to NCBI.

The closest paralog of the gene *HvAs*, the barley high-confidence gene *MLOC_69238* was identified by BLASTn and BLASTp analysis, is allocated to chromosome 6H and was annotated by help of the barley FLcDNA NIASHv1057H10 (GenBank accession ID AK357603). Using *HvAs* and its paralog *MLOC_69238* as query for searching NCBI Protein Database and Phytozyome Database, 48 homologous genes in 20 monocot and dicot species could be determined (Table 3-10). A maximum likelihood-based phylogenetic tree was constructed for these 50 gene sequences (Figure 3-9). The main branch of the tree separated sequences coming from either monocot or dicot species. Within plant families the paralogs formed independent clades if gene sequences were present from different species

(e.g. the *Poaceae*). This strong separation of the paralogous clades in the *Poaceae* may indicate that these paralogs originated by the shared whole-genome duplication that was dated to have occurred around 20 million years before the divergence of the grass species (Paterson et al., 2004; Thiel et al., 2009). The rice orthologs of barley *HvAs* and *MLOC_69238* are located on rice chromosomes *Os*02 and *Os*06, respectively, which represent at least partially orthologous linkage groups to barley chromosomes 6H and 7H (Thiel et al., 2009). This pattern of conserved presence of a pair of ancient segmentally duplicated paralogous genes was found for all diploid species in the phylogenetic analysis as supported by the WGD in the respective clades. Tetraploid species like *Panicum virgatum* and *Brassica rapa* contained two pairs of paralogs indicating that all orthologs of *HvAs* and *MLOC_69238* might have retained an important function in plants.

Table 3-10: Orthologs and segmental paralogs of *HvAs* in monocots and dicots.

Protein ID[A]	Species	Family	cTP Score[B]	CCT Domain[C]	Amino Acid
XP_002864503.1	*Arabidopsis lyrata*	Dicots	0.929	374 - 417	426
XP_002867585.1	*Arabidopsis lyrata*	Dicots	0.970	339 - 382	392
NP_568852.2	*Arabidopsis thaliana*	Dicots	0.920	383 - 426	435
NP_567737.1	*Arabidopsis thaliana*	Dicots	0.979	341 - 384	394
XP_009126853.1	*Brassica rapa*	Dicots	0.968	342 - 385	394
XP_009139739.1	*Brassica rapa*	Dicots	0.893	327 - 372	382
XP_009137564.1	*Brassica rapa*	Dicots	0.928	330 - 375	383
XP_009120241.1	*Brassica rapa*	Dicots	0.970	345 - 387	393
XP_006280489.1	*Capsella rubella*	Dicots	0.856	384 - 427	436
XP_006283882.1	*Capsella rubella*	Dicots	0.966	339 - 382	392
XP_006401232.1	*Eutrema salsugineum*	Dicots	0.646	400 - 443	452
XP_006413244.1	*Eutrema salsugineum*	Dicots	0.913	361 - 406	416
XP_004305701.1	*Fragaria vesca*	Dicots	0.771	378 - 421	433
XP_004296491.1	*Fragaria vesca*	Dicots	0.769	375 - 417	650
XP_003543009.1	*Glycine max*	Dicots	0.852	349 - 391	399
XP_003549301.1	*Glycine max*	Dicots	0.893	347 - 389	398
XP_006597800.1	*Glycine max*	Dicots	0.969	-	395
XP_006586995.1	*Glycine max*	Dicots	0.953	-	374
XP_003610054.1	*Medicago truncatula*	Dicots	0.904	342 - 379	379
XP_003595074.1	*Medicago truncatula*	Dicots	0.987	-	313
EYU38741.1	*Mimulus guttatus*	Dicots	0.590	333 - 375	385
EYU18216.1	*Mimulus guttatus*	Dicots	0.640	362 - 404	410
XP_007154713.1	*Phaseolus vulgaris*	Dicots	0.773	340 - 382	391
XP_007138773.1	*Phaseolus vulgaris*	Dicots	0.353	-	264
XP_002308112.2	*Populus trichocarpa*	Dicots	0.876	392 - 435	448
XP_002324691.2	*Populus trichocarpa*	Dicots	0.807	405 - 448	467
XP_008363941.1	*Malus domestica*	Dicots	0.819	383 - 426	438
XP_008388630.1	*Malus domestica*	Dicots	0.656	381 - 424	436
XP_003560511.1	*Brachypodium distachyon*	Monocots	0.887	394 - 437	448
XP_003570922.1	*Brachypodium distachyon*	Monocots	0.847	380 - 423	434
MLOC_670_HvAs	*Hordeum vulgare*	Monocots	0.742	403 - 448	459
MLOC_69238	*Hordeum vulgare*	Monocots	0.865	436 - 479	490
ABF00989.1	*Oryza sativa*	Monocots	0.928	428 - 471	482
NP_001058473.1	*Oryza sativa*	Monocots	0.790	410 - 455	466
Pavir00048173m	*Panicum virgatum*	Monocots	0.837	-	388
Pavir00001278m	*Panicum virgatum*	Monocots	0.792	-	467
Pavir00019674m	*Panicum virgatum*	Monocots	0.936	-	339
Pavir00054949m	*Panicum virgatum*	Monocots	0.936	-	351
XP_006657299.1	*Oryza brachyantha*	Monocots	0.647	377 - 422	438
XP_006646882.1	*Oryza brachyantha*	Monocots	0.931	426 - 469	480
XP_008811450.1	*Phoenix dactylifera*	Monocots	0.890	360 - 405	436
XP_008799492.1	*Phoenix dactylifera*	Monocots	0.926	356 - 399	411
Si017125m	*Setaria italica*	Monocots	0.922	411 - 454	465
XP_004966116.1	*Setaria italica*	Monocots	0.812	-	459
XP_002451540.1	*Sorghum bicolor*	Monocots	0.927	432 - 475	486
XP_002437531.1	*Sorghum bicolor*	Monocots	0.905	-	457
XP_008644188.1	*Zea mays*	Monocots	0.899	-	435
XP_008645243.1	*Zea mays*	Monocots	0.905	431 - 474	485
AFW75877.1	*Zea mays*	Monocots	0.949	-	443
NP_001147813.1	*Zea mays*	Monocots	0.939	431 - 474	485

[A]- Protein ID coordinated to the accession number on NCBI or Phytozome.

[B]- cTP score was predicted according to the online tool TargetP. It reflects the reliability of chloroplast localization.

[C]- The location of CCT domain was predicted according to InterPro. '-' means no prediction result, i.e. lacking CCT domain.

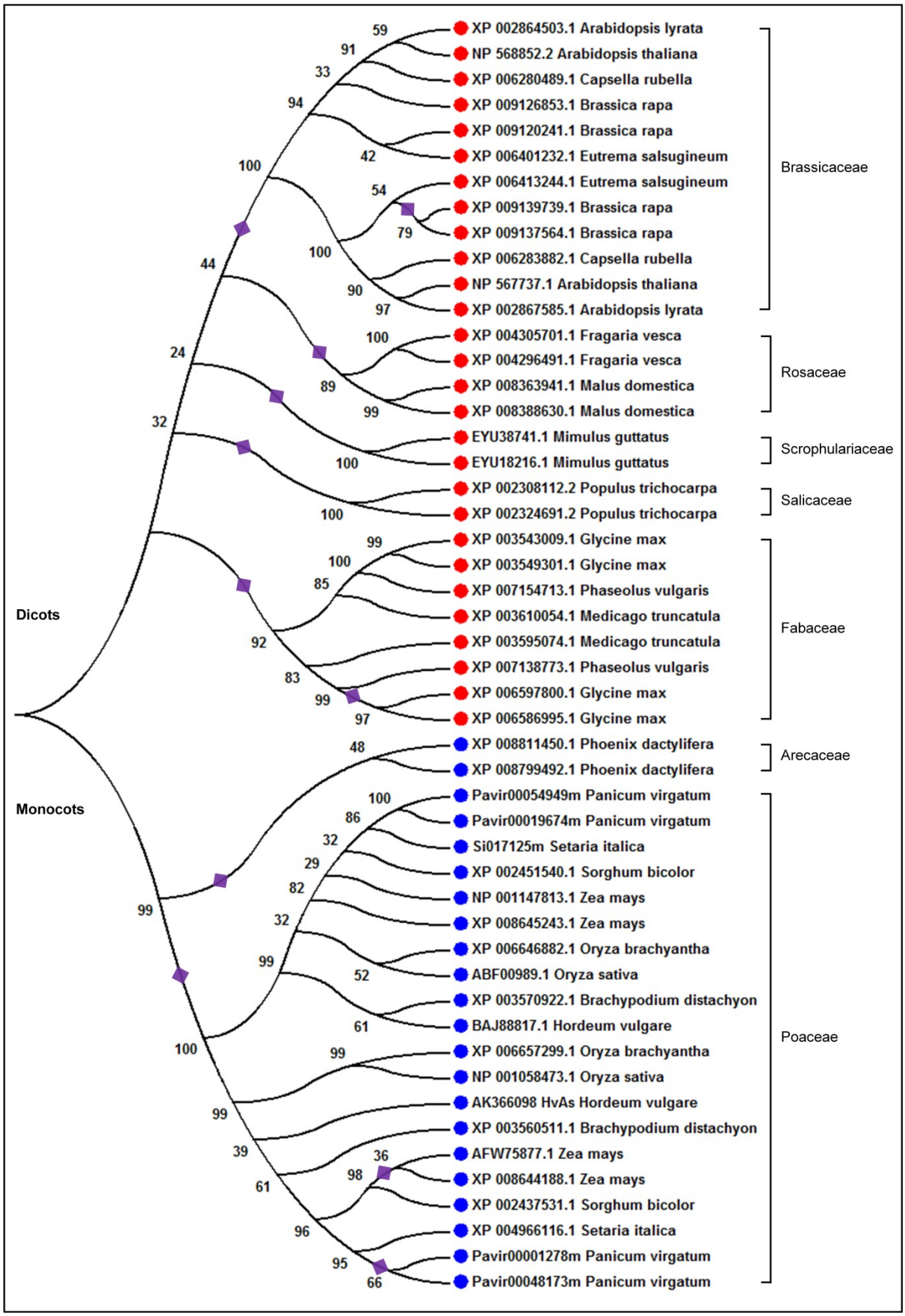

Figure 3-9: Phylogenetic analysis of orthologs and segmental paralogs of the gene *HvAs* from monocot and dicot species. The phylogenetic tree is divided into two main branches separating

monocot and dicot species. In the family of *Poaceae*, orthologs and paralogs of *HvAs* form deeply rooted branches indicating the origin of the paralogous pairs from a WGD event predating the divergence of the *Poaceae* species but occurring after the divergence of *Poaceae* from the *Arecaceae*. A similar pattern as observed for the *Poaceae* is also found for the families of *Brassicaceae* and *Fabaceae* of the dicot branch pointing at WGD events that occured after divergence of the individual family progenitors. The purple squares indicate the WGD events according to Muhlhausen and Kollmar (2013).

4 Discussion

The *albostrians* mutant of barley is one of the most prominent examples used in studies to generate a better understanding of mechanisms underlying chloroplast biogenesis. Over the last three decades research related to the *albostrians* mutant was mainly focused on aspects of molecular biology and biochemistry. A list of biological questions could be addressed by investigating the *albostrians* mutant. For example, studying *albostrians* mutant of barley provided the first evidence for a plastid to nucleus signaling pathway (retrograde signaling, Bradbeer et al., 1979), and it revealed the presence of two RNA polymerase systems in the chloroplast [plastid-encoded plastid RNA polymerase (PEP), (Serino and Maliga, 1998); nucleus-encoded plastid RNA polymerase (NEP), (Liere and Börner, 2011)]. Deeper functional analyses, however, were hampered so far due to the fact that the gene underlying the *albostrians* phenomenon was still unknown. Therefore, the aim of this project was to isolate the gene *albostrians* which was eventually achieved by classical map-based cloning. The identified candidate gene was functionally verified by screening for independent induced alleles in a TILLING population of barley. Functional annotation revealed that the gene *albostrians* represents a member of the CCT motif gene family. Subcellular localization of the ALBOSTRIANS protein provided first evidence of targeting to plastids and maybe also the nucleus (possible dual targeting) which may hint at a role of the protein in regulating gene expression in both compartments. Alternatively, the dual targeting may represent its direct involvement into the retrograde signaling pathway of the plant.

In the current study, identification of the gene *albostrians* benefited from the greatly improved genomic resources of barley as well as *de novo* generated sequence information. Thus the initial mapping, high resolution mapping, physical mapping as well as initial functional validation of the gene could be achieved within the duration of this project.

4.1 The *albostrians* gene of barley is a key factor of chloroplast biogenesis

In this study, the gene *albostrians* (*HvAs*) could be identified by a map-based cloning approach and subsequently, verified by identifying an independent mutant from a barley TILLING population. The functional annotation of the identified gene provided a strong hint that the gene *HvAs* is involved (directly and/or indirectly) in chloroplast

biogenesis. *HvAs* is a putative ortholog of the *chloroplast import apparatus 2* (*CIA2*) gene of *Arabidopsis thaliana* (Sun et al., 2001). Loss-of-function mutations of *AtCIA2* (*Atcia2*) were reported to exhibit a pale green phenotype, which is very distinct to the *albostrians* phenotype. Although in both plant systems mutations in the *CIA2* orthologs affect the chloroplasts, the functional defect still leads to clearly distinct phenotypic characteristics which may be an indication of functional differences of the two genes in Arabidopsis and barley or of differences in the downstream regulatory pathways. CIA2 of Arabidopsis serves as an assisting factor to coordinate protein import into the chloroplasts and protein synthesis within the chloroplasts (Sun et al., 2001; Sun et al., 2009). It belongs to the *CMF* gene family and is considered to act as a transcription factor. The promotor sequences of two classes of genes, encoding plastid localized ribosomal proteins and translocon proteins, respectively, are the main targets of the Arabidopsis CIA2 protein. Other *CMF* gene family members encode transcription factors that modify floral gene expression through DNA-binding or DNA-binding complexes, mediated by the CCT motif (Cockram et al., 2012). Based on its homology to *AtCIA2* the barley gene *HvAs* might as well play an important role in regulating the chloroplast protein import capacity and/or the translation efficiency through regulating the translocon-encoding and ribosomal protein-encoding genes (Figure 4-1B). *In silico* analyses revealed up to 78 putative nucleus-encoded transcription factors localized to chloroplast and with the potential function to regulate expression of chloroplast genes (Wagner and Pfannschmidt, 2006; Schwacke et al., 2007; Liere et al., 2011). Subcellular localization experiments revealed that ALBOSTRIANS is targeted also to the plastid indicating a potential involvement in coordinating plastome gene expression (Figure 4-1A).

Homozygous mutants derived from the M_2-TILLING family 6460-1 exhibited a pure *albino* phenotype in contrast to the three phenotypic classes (*albino*, variegated or green) for which the progeny of a plant carrying the homozygous *albostrians* allele (*as*/*as*) would segregate. The reason for this difference is unclear, however, it may be explained by the possibility that the longer protein translated from the M4205 allele could exhibit residual activity for its essential function. Recently, studies on the *GERANYLGERANYL DIPHOSPHATE SYNTHASE 1* (*GGPS1*) gene of Arabidopsis provided an example how mutations affecting different domains of a gene could result in distinct phenotypes (Ruppel et al., 2013). GGPS1 is a biosynthetic enzyme

in the isoprenoid biosynthetic pathway with the capability of synthesizing geranylgeranyl diphosphate (Okada et al., 2000). The isoprenoid biosynthesis plays an importance role during the early steps of the methylerythritol 4-phosphate (MEP) pathway (Banerjee and Sharkey, 2014). Mutation of genes coding for plastid-localized MEP pathway enzymes required for dimethylallyl diphosphate synthesis can lead to an *albino* phenotype (Mandel et al., 1996; Budziszewski et al., 2001; Gutierrez-Nava Mde et al., 2004; Guevara-Garcia et al., 2005; Hsieh and Goodman, 2006; Hsieh et al., 2008). Whereas, the T-DNA insertion mutant alleles *ggps1-2 and ggps1-3* (insertion occurred after the 42nd and 349th amino acid, respectively) showed a seedling-lethal *albino* phenotype, a third non-synonymous single AA exchange mutant allele *ggps1-1* exhibited a variegated phenotype – white centers with green periphery in the first few rosettle leaves, and new leaves harbor yellowish centers with green periphery in the later development stages (Ruppel et al., 2013). Interpreting the results obtained for different *ggsp1* mutants in Arabidopsis in regard of the two *HvAs* alleles putatively leading to the synthesis of two more or less severely truncated ALBOSTRIANS proteins could imply that the protein resulting from the original M4205 *HvAs* allele exhibits residual activity that is sufficient to trigger normal chloroplast differentiation depending on the developmental circumstances. Since the white sectors have abnormal plastids which are devoid of 70S ribosomes, it might be speculated that proplastids differentiate into mature chloroplasts depending on whether the amount of nucleus-encoded ribosomal proteins imported into the plastids would reach a critical (‘threshold’) level (Figure 4-1C). In case all cells contain plastids with normal sets of ribosomes it would generate pure green leaf. In contrast, cells with ribosome-free plastids give rise to an irreversible *albino* phenotype. The variegated phenotype is formed when neighboring cells contain functional and defective plastids, respectively, which harbor above-threshold amounts of ribosomes and below-threshold ribosomes, respectively.

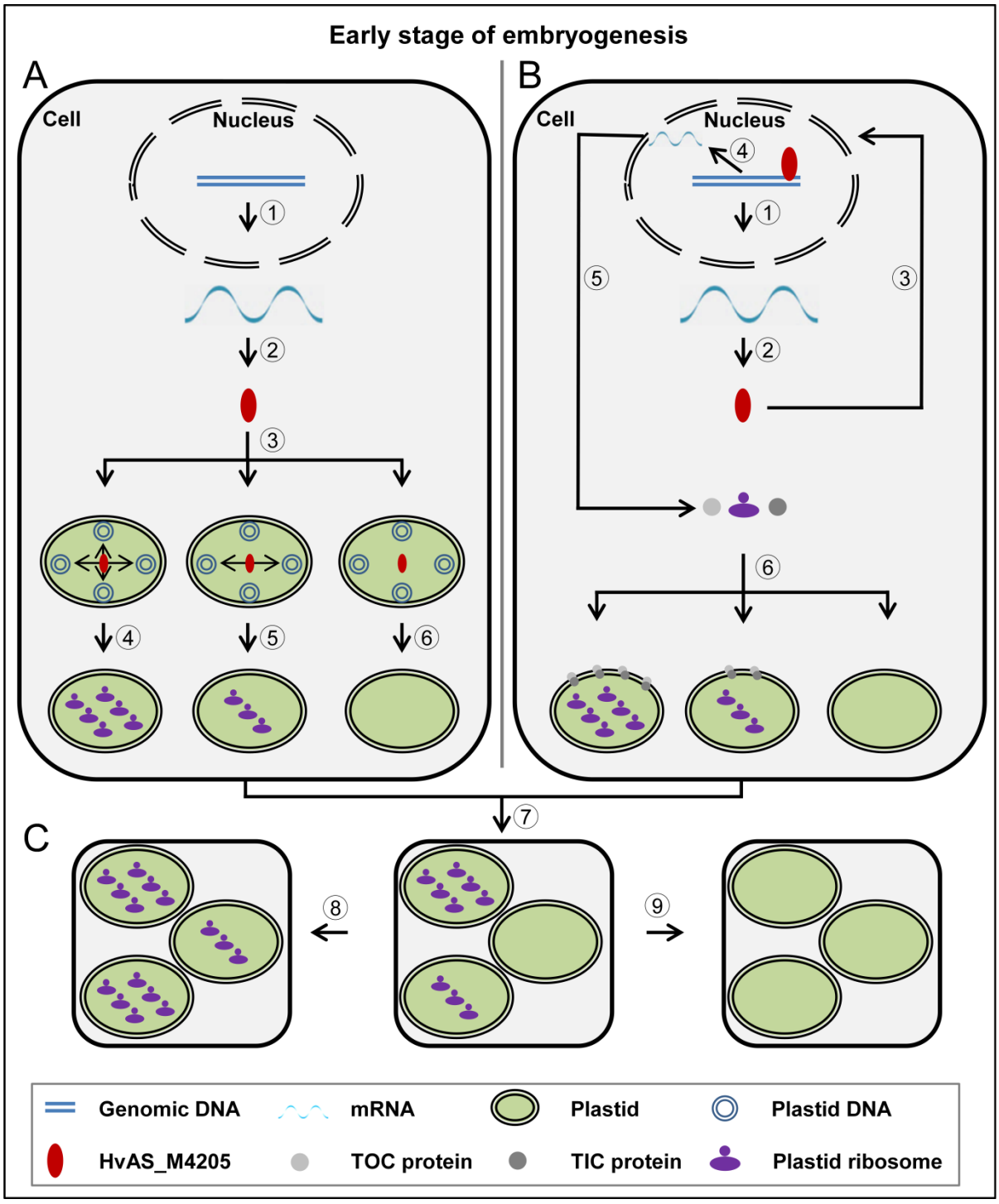

Figure 4-1: Genetic model to explain how the gene *albostrians* could affect the development of chloroplasts. (A) The ALBOSTRIANS protein might be directly involved in gene expression of the plastome. The ALBOSTRIANS protein is considered as an essential factor controlling expression of plastid genes that encode components of plastid 70S ribosomes. It was supposed that the amount of ribosomes in plastids is depending on the efficiency of HvAS_M4205 (section A, step 4-6). (B) Or HvAS_M4205 might play a role as a transcription factor for the expression of genes like *Toc*, *Tic* and *Rib* genes, which are essential for importing ribosomal proteins as well as other nucleus-encoded proteins to the plastids or ribosome synthesis in the cytoplasm. The number of ribosomes in the

plastids is depending on the transport efficiency (section B, step 6). (C) According to the amount of ribosome proteins in the plastids, a threshold model is proposed for the *albostrians* variegation. Cell divisions (step 7) and sorting out of plastids will give rise to healthy cells containing plastids with above-threshold amount of ribosomal proteins (step 8) as well as cells with ribosome-free plastids (step 9). As a consequence, it will form the green and white sectors of the variegated leaf.

4.2 The ALBOSTRIANS protein showed dual localization to chloroplast and nucleus of the same cell in barley

Based on *in silico* analysis a chloroplast transit peptide (cTP) was predicted for the N-terminus of the ALBOSTRIANS protein, implying that the protein would most likely be localized to the chloroplast. The putative protein, however, was also predicted to contain the CCT motif, which was considered before as a nuclear localization signal (Robert et al., 1998). Subcellular localization experiments based on transient expression of ALBOSTRIANS::GFP fusion protein indicated its presence in proplastids and maybe also the nucleus of barley epidermal as well as stomata companion cells. The plastid localization of the fusion protein was confirmed by co-localization of the red-fluorescence protein (mCherry) labeled chloroplast marker pt-rk-CD3-999 (Nelson et al., 2007) whereas the nuclear localization was only supported by strong green fluorescence of the nucleus. However, it remains to be confirmed by independent detection methods like immuno-labeling with ALBOSTRIANS specific antibodies, since unspecific nuclear targeting of GFP was also reported in other systems before (Seibel et al., 2007).

Small and colleagues (1998) postulated the existence of proteins that can be targeted to more than one DNA-containing compartment. Since then, a number of proteins have been identified with dual or multiple localization characteristics (Krause and Krupinska, 2009; Krause et al., 2012). Furthermore, Krause and Krupinska (2009) classified the dual localization proteins into two groups according to their movement pathway. Group I represents *de novo* synthesized proteins with a true dual targeting behavior, i.e. the protein moves to distinct destinations in parallel. DNA LIGASE I of Arabidopsis (Sunderland et al., 2006) and the CARROT DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE (DHFR) (Luo et al., 1997) represent such Group I dual targeted proteins. Group II proteins, in contrast, relocate from organelles to nuclei; however, their primary target is the organelle. So far, four members could be classified to belong to this group, including the plant-specific transcription factor IIB-RELATED PROTEIN pBrp of Arabidopsis (Lagrange

et al., 2003), the TOBACCO STRESS-INDUCED GENE 1 (*Tsi1*) BINDING PROTEIN (TSIP1) (Ham et al., 2006), the transcription factor WHIRLY1 (Krause et al., 2005; Isemer et al., 2012) and the PLANT HOMEODOMAIN (PHD) TRANSCRIPTION FACTOR PTM (for PHD type transcription factor with transmembrane domains) of Arabidopsis (Sun et al., 2011). WHIRLY1 of barley was the first protein to be identified in the nucleus and plastids of the same plant cell (Grabowski et al., 2008). Moreover, a chloroplast transformation approach showed that the AtWHIRLY1 protein was trafficking in a retrograde translocation pathway, i.e. the precursor protein from the cytoplasm was firstly imported into the chloroplast and subsequently relocated into the nucleus after cleavage of the N-terminal plastid transit peptides (Krause and Krupinska, 2009; Isemer et al., 2012). Thus, this novel form of dual targeting involved retrograde translocation from the primary target compartment.

Similar to the above mentioned group II retrograde translocation proteins, the ALBOSTRIANS protein was considered as having transcription factor activity since it contains a CCT motif (Robert et al., 1998). Therefore, based on its potential dual targeting to plastids and nuclei, it might be speculated that the ALBOSTRIANS protein serves as a trans-acting element involved in regulating gene expression in the chloroplast and the nucleus. WHIRLY1 is a member of single-stranded DNA-binding protein gene family and fulfils various functions including maintenance of telomere homeostasis (Yoo et al., 2007), coordination of expression of genes involved in pathogen defense reactions (Desveaux et al., 2000; Xiong et al., 2009) and to maintain plastid genome stability by protecting against repeat-mediated illegitimate recombination (Cappadocia et al., 2010). Regulation of gene expression may occur at different levels involving transcription initiation, transcript processing like splicing and editing as well as translation. Dually targeted proteins with transcription factor properties thus principally bear the potential of being involved in one or the other step of this complex process in the different compartments.

Except for the WHIRLY1 protein (Grabowski et al., 2008), the subcellular distribution of dual-targeted plant proteins has not been investigated by immunohistochemical methods that would require the availability of a highly specific antibody. Further characterization of the ALBOSTRIANS protein by immunological analysis is also required in future studies to resolve better the role of the protein and the processes it might be involved in the different compartments. On the other hand, the localization

pathway of the ALBOSTRIANS protein still remains elusive; it is unclear if the targeting to nucleus and chloroplast is a simultaneous or sequential process. As demonstrated by the example of AtWHIRLY1, chloroplast transformation (van Bel et al., 2001) would provide a possibility to investigate how the ALBOSTRIANS protein moves in the plant cell and thus would contribute to a better understanding of the cross communication processes between chloroplast and the nucleus in the plant cell. Overall, additional future experiments will be required to reproduce and confirm the results of ALBOSTRIANS subcellular localization by alternative experimental approaches. This is especially important, since subcellular co-localization experiments using the two mutated alleles encoding for shortened / truncated ALBOSTRIANS proteins revealed the same dual localization pattern as for the WT protein. This may represent a conflicting result since the nuclear localization signal, the CCT domain (Cockram et al., 2012), was completely eliminated in the two truncated forms, respectively. In case the CCT motif of the ALBOSTRIANS protein has no direct link with nuclear localization, then similarly to WHIRLY1, the ALBOSTRIANS protein may first move to the plastid and then relocate after cleavage of the cTP via the retrograde translocation pathway to the nucleus. This explanation, however, will require further experiments in the future to determine the alternative nuclear localization signal.

4.3 Barley mutants are an important tool for dissecting barley biology and agronomic traits

Mutations can occur spontaneously or may be induced experimentally. Spontaneous mutations form the basis of natural selection and evolution. Artificial mutagenesis is normally achieved by either irradiation or chemical treatment. Since the Swedish geneticists Hermann Nilsson-Ehle and Åke Gustafsson first induced mutations by means of X-rays and UV-irradiations in barley cultivar ‘Gull’ in 1928, a large number of mutants were derived with a broad spectrum of distinct phenotypes (Lundqvist, 2009). These mutant resources provide an important tool for dissecting the genetic basis of important agronomic traits and biological properties of barley. So far, a larger number of barley genes, underlying various key molecular functions, were cloned by help of multiple mutant alleles (Table 4-1). Likewise, in this study, the gene *albostrians* was isolated by using the historical mutant line M4205, derived from X-ray irradiation, as pollen donor for mapping population construction. Subsequently,

screening an EMS induced TILLING population (Gottwald et al., 2009) provided an independent mutated allele which confirmed the function of the identified candidate gene. In this context, further dissection of genetic factors involved in chloroplast biogenesis could benefit from utilizing mutants from historical collections and/or from TILLING populations.

Table 4-1: Genes isolated by taking advantage of barley mutant resources.

Gene ID	Locus name or phenotype	Chromosome Location	Reference
Knox3	hooded phenotype	4H	(Muller et al., 1995)
mlo	mildew resistance locus o	4HL	(Buschges et al., 1997)
rar1	Powdery mildew resistance	2HL	(Shirasu et al., 1999)
Mla	Powdery mildew resistance	1HS	(Halterman et al., 2001)
sln1	slender1	4HS	(Chandler et al., 2002)
Rpg1	Stem rust resistance	7HS	(Brueggeman et al., 2002)
ROR2	Powdery mildew resistance	5HL	(Collins et al., 2003)
lla	shrunken grain	7H	(Morell et al., 2003)
uzu1	semidwarf	3HL	(Chono et al., 2003)
Ppd-H1	Photoperiod-H1	2HS	(Turner et al., 2005)
rym4/rym5	eukaryotic translation initiation factor 4E	3HL	(Stein et al., 2005)
HvEXPB1	root hairless	2H	(Kwasniewski and Szarejko, 2006)
nec1	disease lesion phenotype	1H	(Rostoks et al., 2006)
VRN3	Vernalization 3	7HS	(Yan et al., 2006)
Vrs1	six-rowed spike 1	2H	(Komatsuda et al., 2007)
Nud	naked caryopsis	7HL	(Taketa et al., 2008)
Rpg5/rpg4	Stem rust resistance	5HL	(Brueggeman et al., 2008)
Nec.S1	necrotic steptoe	3H	(Zhang et al., 2009a)
sdw1/denso	semidwarf	3HL	(Jia et al., 2009)
cly1	cleistogamy 1	2HL	(Nair et al., 2010)
lnt1	low number of tillers-1	3HL	(Dabbert et al., 2010)
bgl	betaglucanless	7H	(Taketa et al., 2012)
Lks2	short awn 2	7HL	(Yuo et al., 2012)
Mat.a	early maturity	1HL	(Zakhrabekova et al., 2012)
Trd1	third outer glume 1	1HL	(Houston et al., 2012)
Vrs4	six-rowed spike 4	3HS	(Koppolu et al., 2013)
mnd	many-noded dwarf	5HL	(Mascher et al., 2014)

4.4 Improved genomic resources greatly facilitate gene isolation in barley

In the *Triticeae* crops like barley with multi-gigabase genomes and lacking of a reference genome sequence, forward genetic gene isolation was a time consuming and labor-intensive endeavor. Especially the procedure of marker development and chromosome walking for establishing a local physical map at the locus of interest

created major challenges and for long made it impossible to accomplish genetic mapping, gene isolation and functional characterization of a candidate gene within the duration of a single thesis. In the present study, cloning of the gene *albostrians* took advantage of the massively improved barley genomic resources. Initially, the gene *albostrians* was allocated to barley chromosome 7H on the basis of 381 BOPA markers selected from a publicly available resource comprising 2,943 gene-based SNP markers (Close et al., 2009). Subsequently, saturation mapping the target interval was relying further on publicly available genomic resources (Sato et al., 2009a; Sato et al., 2009b; Mayer et al., 2011; International Barley Genome Sequencing Consortium, 2012), which, mainly on the basis of *in silico* analyses, could deliver the design of 63 polymorphic markers. As a result, the gene *albostrians* was allocated quickly to a genetic interval of around 0.06 cM. No chromosome walking was required since the two flanking markers Zip_2661 and Zip_2665 could be anchored to the barley physical map and sequence resource (International Barley Genome Sequencing Consortium, 2012; Mascher et al., 2013a) by simple sequence comparison (Altschul et al., 1990). Immediate sequencing of 60 MTP BAC clones representing the identified five physical map contigs allowed to establish a contiguous sequence scaffold and delimiting flanking markers and the *albostrians* candidate gene to a physical distance of about 46 Kbp. Overall this procedure is very much facilitated now and the challenges are comparable to work in model species like Arabidopsis and rice. The remaining species-specific time-limiting steps are imposed by the generation time of barley and still existing limitations in gene transformation-dependent functional confirmation strategies (Harwood, 2012), which complicated the experimental validation of the *albostrians* candidate gene by transgene analysis.

Future attempts of forward genetic gene isolation in barley may profit furthermore from innovative approaches introduced on the basis of Next Generation Sequencing (NGS) technology. The lower cost of high throughput next generation sequencing provides the possibility of sequencing individual plants to perform genome-wide SNP discovery (Kircher and Kelso, 2010). Studies in peach (Ahmad et al., 2011) and barley (International Barley Genome Sequencing Consortium, 2012) proved that abundant high quality SNPs derived from whole genome shotgun sequencing could be directly used for genetic mapping and diversity studies. This procedure was also

already applied to the advantage of the cloning of the gene *albostrians*, when survey sequencing of the mutant plant M4205 revealed 550,000 and 92,000 SNP at whole genome or chromosome 7H scale, respectively, between the two parental genotypes Morex and M4205 of the mapping population. Besides M4205, also the original wild type genotype Haisa was included for survey sequencing and 65,000 SNPs were identified in comparison to cultivar Morex. This marker resource may be directly utilized in the future to exploit the 'Scholz mutant collection' at IPK. This collection was the source of the *albostrians* mutant M4205 and contains at least 76 mutant lines with genetic background of Haisa (communicated with Dr. Helmut Knüpffer, Head of AG Genebank Documentation, IPK), thus future mapping of the underlying genes can take advantage of the established SNP resource.

In comparison to the whole genome shotgun sequencing strategy, methods like RNA-seq (James et al., 2013; Ramirez-Gonzalez et al., 2014; Schneeberger, 2014), genotyping-by-sequencing (GBS) (Poland et al., 2012) and exome capture based re-sequencing (Mascher et al., 2013b) introduce levels of complexity reduction thus open the possibility for cost-efficient high-throughput marker development or even for direct gene identification. As an example, the *many-noded dwarf* (*mnd*) mutant of barley could be recently cloned through an exome capture based mapping-by-sequencing strategy (Mascher et al., 2014). Altogether, the NGS technologies are expected to advance barley genetics and crop improvement in general.

5 Outlook

Identification of nuclear genes controlling variegated phenotype in plants is an essential prerequisite towards studying and understanding basic mechanisms of chloroplast biogenesis. In this context, three biological questions need to be addressed: Which gene is responsible for the variegated phenotype? How does the gene product act at the cellular level? What is the difference in fulfilling this activity in wild type (green) compared to mutant (*albino*) tissue?

The current study focussed on the initial question, to isolate the barley gene conferring the *albostrians* specific variegated phenotype. Through a map-based cloning approach, the gene *albostrians* was cloned and subsequently confirmed by screening a TILLING population providing evidence that the identified *albostrians* candidate gene is needed for the development of green, photosynthetic active chloroplast. Preliminary subcellular co-localization experiments indicated that the ALBOSTRIANS protein may be directed to both nucleus and chloroplast. Whether this dual localization is connected to a functional role of the protein in both cellular compartments or if one compartment acts like a buffer or depot of the (inactive) form of the factor requires further experiments in the future. Besides gene function confirmation by mutant analysis, it will be important to generate an independent functional evidence for the gene function. Here, a transgenic approach, e.g. complementation of the mutant M4205 by expressing the WT *Albostrians* gene provides an attractive option; however, success of this approach is very much dependent on the possibility to regenerate genetically transformed plants from the M4205 genotype. To further support the subcellular localization results immunolabeling of the protein and detection by Transmission Electron Microscopy (TEM) would provide an independent alternative approach.

Functional annotation of the ALBOSTRIANS protein revealed it to belong to the *CMF* gene family. Members of this gene family are potentially involved in DNA-binding or forming protein-protein complexes (Cockram et al., 2012). If HvAS is thus a transcriptional regulator it would be promising to perform a chromatin immunoprecipitation sequencing (ChIP-Seq) experiment (Saleh et al., 2008; Valouev et al., 2008) in order to reveal target genes or a target gene network that is regulated by the *HvAs* gene. As demonstrated by the *IDEAL PLANT ARCHITECTURE1* (*IPA1*)

gene in rice (Lu et al., 2013), integration of genomic data can be used to reveal a complex regulating gene network involved in chloroplast biogenesis. This can be substantiated by deep transcriptome sequencing (RNA-Seq), however, such attempts need to consider any masking or pleiotropic effects if comparing green *vs*. white tissue, since the blocking of chloroplast biogenesis will very likely affect many if not all regulatory pathways of a cell. To elucidate if HvAS acts as a monomer or in homo- or hetero-multimers, the use of yeast two-hybrid (Y2H) interaction screening (Cao and Yan, 2013) would be suitable to identify potential interacting protein partners. This would help to characterize function of the ALBOSTRIANS protein and provide evidence for the next steps towards identification of novel genetic factors involved in chloroplast biogenesis.

6 References

Abdelnoor, R.V., Yule, R., Elo, A., Christensen, A.C., Meyer-Gauen, G., and Mackenzie, S.A. (2003). Substoichiometric shifting in the plant mitochondrial genome is influenced by a gene homologous to MutS. Proc Natl Acad Sci U S A **100,** 5968-5973.

Ahmad, R., Parfitt, D.E., Fass, J., Ogundiwin, E., Dhingra, A., Gradziel, T.M., Lin, D., Joshi, N.A., Martinez-Garcia, P.J., and Crisosto, C.H. (2011). Whole genome sequencing of peach (Prunus persica L.) for SNP identification and selection. BMC Genomics **12,** 569.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of molecular biology **215,** 403-410.

Ankele, E., Kindgren, P., Pesquet, E., and Strand, A. (2007). In vivo visualization of Mg-protoporphyrin IX, a coordinator of photosynthetic gene expression in the nucleus and the chloroplast. Plant Cell **19,** 1964-1979.

Ariyadasa, R., and Stein, N. (2012). Advances in BAC-based physical mapping and map integration strategies in plants. Journal of biomedicine & biotechnology **2012,** 184854.

Ariyadasa, R., Mascher, M., Nussbaumer, T., Schulte, D., Frenkel, Z., Poursarebani, N., Zhou, R., Steuernagel, B., Gundlach, H., Taudien, S., Felder, M., Platzer, M., Himmelbach, A., Schmutzer, T., Hedley, P.E., Muehlbauer, G.J., Scholz, U., Korol, A., Mayer, K.F., Waugh, R., Langridge, P., Graner, A., and Stein, N. (2014). A sequence-ready physical map of barley anchored genetically by two million single-nucleotide polymorphisms. Plant Physiol **164,** 412-423.

Banerjee, A., and Sharkey, T.D. (2014). Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. Natural product reports **31,** 1043-1055.

Bedbrook, J.R., and Bogorad, L. (1976). Endonuclease recognition sites mapped on Zea mays chloroplast DNA. Proc Natl Acad Sci U S A **73,** 4309-4313.

Bertioli, D.J., Ozias-Akins, P., Chu, Y., Dantas, K.M., Santos, S.P., Gouvea, E., Guimaraes, P.M., Leal-Bertioli, S.C., Knapp, S.J., and Moretzsohn, M.C. (2014). The use of SNP markers for linkage mapping in diploid and tetraploid peanuts. G3 **4,** 89-96.

Bogorad, L., and Woodcock, C.L.F. (1970). Autonomy and Biogenesis of Mitochondria and Chloroplasts **In North-Holland, Amsterdam,** 92-97.

Bolser, D.M., Kerhornou, A., Walts, B., and Kersey, P. (2015). Triticeae resources in Ensembl Plants. Plant Cell Physiol **56,** e3.

Börner, T. (1986). Chloroplast control of nuclear gene function. Endocyt. C. Res. **3,** 265-274.

Börner, T., and Sears, B.B. (1986). Plastome mutants. Plant Molecular Biology Reporter **2,** 69-92.

Börner, T., Schumann, B., and Hagemann, R. (1976). Biochemical studies on a plastid ribosome-deficient mutant of *hordeum vulgare*. In Th Bucher, W Neupert, W Sebald, S Werner, (eds)Gentics and Biogenesis of Chloroplast and Mitochondria. Elsevier/North Holland Biomedical Press, Amsterdam**,** pp. 41-48.

Börner, T., Manteuffel, R., and Wellburn, A.R. (1979). Enzymes of Plastid Ribosome-Deficient Mutants Chloroplast ATPase (CF1). Protoplasma **98,** 153-161.

Börner, T., Mendel, R.R., and Schiemann, J. (1986). Nitrate reductase is not accumulated in chloroplast-ribosome-deficient mutants of higher plants. Planta **169,** 202-207.

Bradbeer, J.W., and Börner, T. (1978). Activities of glyceraldehyde-phosphate dehydrogenase ($NADP^+$) and phosphoribulokinase in two barley mutants deficient in chloroplast ribosomes. In G. Akoyunoglou et al. (eds) Chloroplast Development**,** pp. 727-732.

Bradbeer, J.W., Atkinson, Y.E., Börner, T., and HAGEMANN, R. (1979). Cytoplasmic synthesis of plastid polypeptides may be controlled by plastid-synthesised RNA. Nature **279,** 816-817.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., and Kleinhofs, A. (2002). The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. Proc Natl Acad Sci U S A **99,** 9328-9333.

Brueggeman, R., Druka, A., Nirmala, J., Cavileer, T., Drader, T., Rostoks, N., Mirlohi, A., Bennypaul, H., Gill, U., Kudrna, D., Whitelaw, C., Kilian, A., Han, F., Sun, Y., Gill, K., Steffenson, B., and Kleinhofs, A. (2008). The stem rust resistance gene Rpg5 encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains. Proc Natl Acad Sci U S A **105,** 14970-14975.

Budziszewski, G.J., Lewis, S.P., Glover, L.W., Reineke, J., Jones, G., Ziemnik, L.S., Lonowski, J., Nyfeler, B., Aux, G., Zhou, Q., McElver, J., Patton, D.A., Martienssen, R., Grossniklaus, U., Ma, H., Law, M., and Levin, J.Z. (2001). Arabidopsis genes essential for seedling viability: isolation of insertional mutants and molecular cloning. Genetics **159,** 1765-1778.

Buschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Topsch, S., Vos, P., Salamini, F., and Schulze-Lefert, P. (1997). The barley Mlo gene: a novel control element of plant pathogen resistance. Cell **88,** 695-705.

Cao, S., and Yan, L. (2013). Construction of a high-quality yeast two-hybrid (Y2H) library and its application in identification of interacting proteins with key vernalization regulator TaVRN-A1 in wheat. BMC Res Notes **6,** 81.

Cappadocia, L., Marechal, A., Parent, J.S., Lepage, E., Sygusch, J., and Brisson, N. (2010). Crystal structures of DNA-Whirly complexes and their role in Arabidopsis organelle genome repair. Plant Cell **22,** 1849-1867.

Carol, P., Stevenson, D., Bisanz, C., Breitenbach, J., Sandmann, G., Mache, R., Coupland, G., and Kuntz, M. (1999). Mutations in the Arabidopsis gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. Plant Cell **11,** 57-68.

Carollo, V., Matthews, D.E., Lazo, G.R., Blake, T.K., Hummel, D.D., Lui, N., Hane, D.L., and Anderson, O.D. (2005). GrainGenes 2.0. an improved resource for the small-grains community. Plant Physiol **139,** 643-651.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res **39,** e82.

Chandler, P.M., Marion-Poll, A., Ellis, M., and Gubler, F. (2002). Mutants at the Slender1 locus of barley cv Himalaya. Molecular and physiological characterization. Plant Physiol **129,** 181-190.

Chen, M., Choi, Y., Voytas, D.F., and Rodermel, S. (2000). Mutations in the Arabidopsis VAR2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. Plant J **22,** 303-313.

Chen, W., Chen, H., Zheng, T., Yu, R., Terzaghi, W.B., Li, Z., Deng, X.W., Xu, J., and He, H. (2014). Highly efficient genotyping of rice biparental populations by GoldenGate assays based on parental resequencing. Theor Appl Genet **127,** 297-307.

Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H., and Sheen, J. (1996). Engineered GFP as a vital reporter in plants. Curr Biol **6,** 325-330.

Chono, M., Honda, I., Zeniya, H., Yoneyama, K., Saisho, D., Takeda, K., Takatsuto, S., Hoshino, T., and Watanabe, Y. (2003). A semidwarf phenotype of barley uzu results from a nucleotide substitution in the gene encoding a putative brassinosteroid receptor. Plant Physiol **133,** 1209-1219.

Chumley, T.W., Palmer, J.D., Mower, J.P., Fourcade, H.M., Calie, P.J., Boore, J.L., and Jansen, R.K. (2006). The complete chloroplast genome sequence of Pelargonium x hortorum: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol **23,** 2175-2190.

Chun, E.H., Vaughan, M.H., Jr., and Rich, A. (1963). The Isolation and Characterization of DNA Associated with Chloroplast Preparations. Journal of molecular biology **7,** 130-141.

Close, T.J., Wanamaker, S.I., Caldo, R.A., Turner, S.M., Ashlock, D.A., Dickerson, J.A., Wing, R.A., Muehlbauer, G.J., Kleinhofs, A., and Wise,

R.P. (2004). A new resource for cereal genomics: 22K barley GeneChip comes of age. Plant Physiol **134,** 960-968.

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S., Bozdag, S., Roose, M.L., Moscou, M.J., Chao, S., Varshney, R.K., Szucs, P., Sato, K., Hayes, P.M., Matthews, D.E., Kleinhofs, A., Muehlbauer, G.J., DeYoung, J., Marshall, D.F., Madishetty, K., Fenton, R.D., Condamine, P., Graner, A., and Waugh, R. (2009). Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics **10,** 582.

Cockram, J., Thiel, T., Steuernagel, B., Stein, N., Taudien, S., Bailey, P.C., and O'Sullivan, D.M. (2012). Genome dynamics explain the evolution of flowering time CCT domain gene families in the Poaceae. PLoS One **7,** e45307.

Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., and Henikoff, S. (2001). High-throughput screening for induced point mutations. Plant Physiol **126,** 480-484.

Collins, N.C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J.L., Huckelhoven, R., Stein, M., Freialdenhoven, A., Somerville, S.C., and Schulze-Lefert, P. (2003). SNARE-protein-mediated disease resistance at the plant cell wall. Nature **425,** 973-977.

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., Shaw, P., Bayer, M., Thomas, W., Marshall, D., Hedley, P., Tondelli, A., Pecchioni, N., Francia, E., Korzun, V., Walther, A., and Waugh, R. (2012). Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nat Genet **44,** 1388-1392.

Dabbert, T., Okagaki, R.J., Cho, S., Heinen, S., Boddu, J., and Muehlbauer, G.J. (2010). The genetics of barley low-tillering mutants: low number of tillers-1 (lnt1). Theor Appl Genet **121,** 705-715.

de Felipe, P. (2004). Skipping the co-expression problem: the new 2A "CHYSEL" technology. Genetic vaccines and therapy **2,** 13.

de Felipe, P., Luke, G.A., Hughes, L.E., Gani, D., Halpin, C., and Ryan, M.D. (2006). E unum pluribus: multiple proteins from a self-processing polyprotein. Trends in biotechnology **24,** 68-75.

Desveaux, D., Despres, C., Joyeux, A., Subramaniam, R., and Brisson, N. (2000). PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. Plant Cell **12,** 1477-1489.

Deulvot, C., Charrel, H., Marty, A., Jacquin, F., Donnadieu, C., Lejeune-Henaut, I., Burstin, J., and Aubert, G. (2010). Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. BMC Genomics **11,** 468.

Doyle, J.J., and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. Focus **12,** 13-15.

Ellis, R.J. (1977). Biochem. Biophys. Acta **463,** 185-215.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein science : a publication of the Protein Society **8,** 978-984.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. Journal of molecular biology **300,** 1005-1016.

Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res **8,** 186-194.

Ewing, B., Hillier, L., Wendl, M.C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res **8,** 175-185.

Feschotte, C., Jiang, N., and Wessler, S.R. (2002). Plant transposable elements: where genetics meets genomics. Nat Rev Genet **3,** 329-341.

Froger, A., and Hall, J.E. (2007). Transformation of plasmid DNA into E. coli using the heat shock method. Journal of visualized experiments : JoVE**,** 253.

Gawronski, P., Ariyadasa, R., Himmelbach, A., Poursarebani, N., Kilian, B., Stein, N., Steuernagel, B., Hensel, G., Kumlehn, J., Sehgal, S.K., Gill, B.S., Gould, P., Hall, A., and Schnurbusch, T. (2014). A distorted circadian clock causes early flowering and temperature-dependent variation in spike development in the Eps-3Am mutant of einkorn wheat. Genetics **196,** 1253-1261.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res **40,** D1178-1186.

Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U., and Stein, N. (2009). TILLING in the two-rowed barley cultivar 'Barke' reveals preferred sites of functional diversity in the gene HvHox1. BMC Res Notes **2,** 258.

Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. (2008). Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. Plant Physiol **147,** 1800-1804.

Guevara-Garcia, A., San Roman, C., Arroyo, A., Cortes, M.E., de la Luz Gutierrez-Nava, M., and Leon, P. (2005). Characterization of the Arabidopsis clb6 mutant illustrates the importance of posttranscriptional regulation of the methyl-D-erythritol 4-phosphate pathway. Plant Cell **17,** 628-643.

Guisinger, M.M., Kuehl, J.V., Boore, J.L., and Jansen, R.K. (2011). Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. Mol Biol Evol **28,** 583-600.

Gurushidze, M., Hensel, G., Hiekel, S., Schedel, S., Valkov, V., and Kumlehn, J. (2014). True-Breeding Targeted Gene Knock-Out in Barley Using Designer TALE-Nuclease in Haploid Cells. PLoS One **9,** e92046.

Gutierrez-Nava Mde, L., Gillmor, C.S., Jimenez, L.F., Guevara-Garcia, A., and Leon, P. (2004). CHLOROPLAST BIOGENESIS genes act cell and noncell autonomously in early chloroplast development. Plant Physiol **135,** 471-482.

Hagemann, R., and Scholz, F. (1962). A case of gene induced mutations of the plasmotype in barley. Theor. Appl. Genet. **32,** 50-59.

Hagemann, R., and Börner, T. (1978). Plastid ribosome-deficient mutants of higher plants as a tool in studying chloroplast biogenesis. In G Akouyunoglou, JH Argyroudi-Akoyunoglou, (eds) Chloroplast Development. Elsevier/North Holland Biomedical Press, Amsterdam**,** pp. 709-720.

Hajdukiewicz, P.T., Allison, L.A., and Maliga, P. (1997). The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J **16,** 4041-4048.

Hall, B.G. (2013). Building phylogenetic trees from molecular data with MEGA. Mol Biol Evol **30,** 1229-1235.

Hall, L.N., Roth, R., Brutnell, T.P., and Langdale, J.A. (1998). Cellular differentiation in the maize leaf is disrupted by bundle sheath defective mutations. Symposia of the Society for Experimental Biology **51,** 27-31.

Halterman, D., Zhou, F., Wei, F., Wise, R.P., and Schulze-Lefert, P. (2001). The MLA6 coiled-coil, NBS-LRR protein confers AvrMla6-dependent resistance specificity to Blumeria graminis f. sp. hordei in barley and wheat. Plant J **25,** 335-348.

Ham, B.K., Park, J.M., Lee, S.B., Kim, M.J., Lee, I.J., Kim, K.J., Kwon, C.S., and Paek, K.H. (2006). Tobacco Tsip1, a DnaJ-type Zn finger protein, is recruited to and potentiates Tsi1-mediated transcriptional activation. Plant Cell **18,** 2005-2020.

Han, C.D., Coe, E.H., Jr., and Martienssen, R.A. (1992). Molecular cloning and characterization of iojap (ij), a pattern striping gene of maize. EMBO J **11,** 4037-4046.

Harwood, W.A. (2012). Advances and remaining challenges in the transformation of barley and wheat. J Exp Bot **63,** 1791-1798.

Hedtke, B., Wagner, I., Borner, T., and Hess, W.R. (1999). Inter-organellar crosstalk in higher plants: impaired chloroplast development affects mitochondrial gene and transcript levels. Plant J **19,** 635-643.

Hensel, G., Valkov, V., Middlefell-Williams, J., and Kumlehn, J. (2008). Efficient generation of transgenic barley: the way forward to modulate plant-microbe interactions. Journal of plant physiology **165,** 71-82.

Hess, W.R., Hübschmann, T., and Börner, T. (1994a). Ribosome-deficient plastids of *albostrians* barley: extreme representatives of non-photosynthetic plalstids. Endocytobiosis & Cell Research **10,** 65-80.

Hess, W.R., Muller, A., Nagy, F., and Borner, T. (1994b). Ribosome-deficient plastids affect transcription of light-induced nuclear genes: genetic evidence for a plastid-derived signal. Mol Gen Genet **242,** 305-312.

Hess, W.R., Schendel, R., Rudiger, W., Fieder, B., and Borner, T. (1992). Components of chlorophyll biosynthesis in a barley albina mutant unable to synthesize delta-aminolevulinic acid by utilizing the transfer RNA for glutamic acid. Planta **188,** 19-27.

Hess, W.R., Prombona, A., Fieder, B., Subramanian, A.R., and Borner, T. (1993). Chloroplast rps15 and the rpoB/C1/C2 gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. EMBO J **12,** 563-571.

Himmelbach, A., Zierold, U., Hensel, G., Riechen, J., Douchkov, D., Schweizer, P., and Kumlehn, J. (2007). A set of modular binary vectors for transformation of cereals. Plant Physiol **145,** 1192-1200.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.R., Meng, B.Y., and et al. (1989). The complete sequence of the rice (Oryza sativa) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet **217,** 185-194.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res **35,** W585-587.

Houston, K., Druka, A., Bonar, N., Macaulay, M., Lundqvist, U., Franckowiak, J., Morgante, M., Stein, N., and Waugh, R. (2012). Analysis of the barley bract suppression gene Trd1. Theor Appl Genet **125,** 33-45.

Hsieh, M.H., and Goodman, H.M. (2005). The Arabidopsis IspH homolog is involved in the plastid nonmevalonate pathway of isoprenoid biosynthesis. Plant Physiol **138,** 641-653.

Hsieh, M.H., and Goodman, H.M. (2006). Functional evidence for the involvement of Arabidopsis IspF homolog in the nonmevalonate pathway of plastid isoprenoid biosynthesis. Planta **223,** 779-784.

Hsieh, M.H., Chang, C.Y., Hsu, S.J., and Chen, J.J. (2008). Chloroplast localization of methylerythritol 4-phosphate pathway enzymes and regulation of mitochondrial genes in ispD and ispE albino mutants in Arabidopsis. Plant Mol Biol **66,** 663-673.

International Barley Genome Sequencing Consortium. (2012). A physical, genetic and functional sequence assembly of the barley genome. Nature **491,** 711-716.

Isemer, R., Mulisch, M., Schafer, A., Kirchner, S., Koop, H.U., and Krupinska, K. (2012). Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. FEBS Lett **586,** 85-88.

James, G.V., Patel, V., Nordstrom, K.J., Klasen, J.R., Salome, P.A., Weigel, D., and Schneeberger, K. (2013). User guide for mapping-by-sequencing in Arabidopsis. Genome biology **14,** R61.

Jensen, P.E., and Leister, D. (2014). Chloroplast evolution, structure and functions. F1000prime reports **6,** 40.

Jia, Q., Zhang, J., Westcott, S., Zhang, X.Q., Bellgard, M., Lance, R., and Li, C. (2009). GA-20 oxidase as a candidate for the semidwarf gene sdw1/denso in barley. Funct Integr Genomics **9,** 255-262.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., and Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics.

Joung, J.K., and Sander, J.D. (2013). TALENs: a widely applicable technology for targeted genome editing. Nature reviews. Molecular cell biology **14,** 49-55.

Keddie, J.S., Carroll, B., Jones, J.D., and Gruissem, W. (1996). The DCL gene of tomato is required for chloroplast development and palisade cell morphogenesis in leaves. EMBO J **15,** 4208-4217.

Kircher, M., and Kelso, J. (2010). High-throughput DNA sequencing--concepts and limitations. Bioessays **32,** 524-536.

Kirk, J.T.O., and Tilney-Bassett, R.A.E. (1978). The Plastids, 2nd edn. Elsevier/North-Holland, Amsterdam, the Netherlands.

Komatsuda, T., Pourkheirandish, M., He, C., Azhaguvel, P., Kanamori, H., Perovic, D., Stein, N., Graner, A., Wicker, T., Tagiri, A., Lundqvist, U., Fujimura, T., Matsuoka, M., Matsumoto, T., and Yano, M. (2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci U S A **104,** 1424-1429.

Koppolu, R., Anwar, N., Sakuma, S., Tagiri, A., Lundqvist, U., Pourkheirandish, M., Rutten, T., Seiler, C., Himmelbach, A., Ariyadasa, R., Youssef, H.M., Stein, N., Sreenivasulu, N., Komatsuda, T., and Schnurbusch, T. (2013). Six-rowed spike4 (Vrs4) controls spikelet determinacy and row-type in barley. Proc Natl Acad Sci U S A **110,** 13198-13203.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics **23,** 1289-1291.

Krattinger, S., Wicker, T., and Keller, B. (2009). Map-based cloning of genes in Triticeae (Wheat and Barley). C. Feuillet, G.J. Muehlbauer (eds.), *Genetics and Genomics of the Triticeae*, Plant Genetics and Genomics: Crops and Models **7,** pp 337–357.

Krause, K., and Krupinska, K. (2009). Nuclear regulators with a second home in organelles. Trends Plant Sci **14,** 194-199.

Krause, K., Oetke, S., and Krupinska, K. (2012). Dual targeting and retrograde translocation: regulators of plant nuclear gene expression can be sequestered by plastids. International journal of molecular sciences **13,** 11085-11101.

Krause, K., Kilbienski, I., Mulisch, M., Rodiger, A., Schafer, A., and Krupinska, K. (2005). DNA-binding proteins of the Whirly family in Arabidopsis thaliana are targeted to the organelles. FEBS Lett **579,** 3707-3712.

Krol, A.R.v.d., Lenting, P.E., Veenstra, J., Meer, I.M.v.d., Koes, R.E., Gerats, A.G.M., Mol, J.N.M., and R.Stuitje, A. (1988). An antisense chalcone synthase gene in transgenic plants inhibits flower pigmentation. Nature **3331,** 866–869.

Krupinska, K., Melonek, J., and Krause, K. (2013). New insights into plastid nucleoid structure and functionality. Planta **237,** 653-664.

Kunzel, G., Korzun, L., and Meister, A. (2000). Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics **154,** 397-412.

Kwasniewski, M., and Szarejko, I. (2006). Molecular cloning and characterization of beta-expansin gene related to root hair formation in barley. Plant Physiol **141,** 1149-1158.

Lagrange, T., Hakimi, M.A., Pontier, D., Courtois, F., Alcaraz, J.P., Grunwald, D., Lam, E., and Lerbs-Mache, S. (2003). Transcription factor IIB (TFIIB)-related protein (pBrp), a plant-specific member of the TFIIB-related protein family. Mol Cell Biol **23,** 3274-3286.

Li, H.M., and Chiu, C.C. (2010). Protein Transport into Chloroplasts. Annual Review of Plant Biology, Vol 61 **61,** 157-180.

Liere, K., and Börner, T. (2011). Transcription in plant mitochondria. F. Kempken (eds) Plant Mitochondria, Advance in Plant Biology

Liere, K., Weihe, A., and Borner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. Journal of plant physiology **168,** 1345-1360.

Liu, X., Yu, F., and Rodermel, S. (2010). An Arabidopsis pentatricopeptide repeat protein, SUPPRESSOR OF VARIEGATION7, is required for FtsH-mediated chloroplast biogenesis. Plant Physiol **154,** 1588-1601.

Lopez-Juez, E., and Pyke, K.A. (2005). Plastids unleashed: their development and their integration in plant development. The International journal of developmental biology **49,** 557-577.

Lu, Z., Yu, H., Xiong, G., Wang, J., Jiao, Y., Liu, G., Jing, Y., Meng, X., Hu, X., Qian, Q., Fu, X., Wang, Y., and Li, J. (2013). Genome-wide binding analysis of the transcription activator ideal plant architecture1 reveals a complex network regulating rice plant architecture. Plant Cell **25,** 3743-3759.

Lundqvist, U. (2009). Eighty Years of Scandinavian Barley Mutation Genetics and Breeding. In QY Shu (eds), Induced Plant Mutations in the Genomics Era**,** 39-43.

Luo, M., Orsi, R., Patrucco, E., Pancaldi, S., and Cella, R. (1997). Multiple transcription start sites of the carrot dihydrofolate reductase-thymidylate synthase gene, and sub-cellular localization of the bifunctional protein. Plant Mol Biol **33,** 709-722.

Maliga, P. (2004). Plastid transformation in higher plants. Annu Rev Plant Biol **55,** 289-313.

Mandel, M.A., Feldmann, K.A., Herrera-Estrella, L., Rocha-Sosa, M., and Leon, P. (1996). CLA1, a novel gene required for chloroplast development, is highly conserved in evolution. Plant J **9,** 649-658.

Marcel, T.C., Varshney, R.K., Barbieri, M., Jafary, H., de Kock, M.J., Graner, A., and Niks, R.E. (2007). A high-density consensus map of barley to compare the distribution of QTLs for partial resistance to Puccinia hordei and of defence gene homologues. Theor Appl Genet **114,** 487-500.

Mascher, M., Jost, M., Kuon, J.E., Himmelbach, A., Assfalg, A., Beier, S., Scholz, U., Graner, A., and Stein, N. (2014). Mapping-by-sequencing accelerates forward genetics in barley. Genome biology **15,** R78.

Mascher, M., Muehlbauer, G.J., Rokhsar, D.S., Chapman, J., Schmutz, J., Barry, K., Munoz-Amatriain, M., Close, T.J., Wise, R.P., Schulman, A.H., Himmelbach, A., Mayer, K.F., Scholz, U., Poland, J.A., Stein, N., and Waugh, R. (2013a). Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). Plant J **76,** 718-727.

Mascher, M., Richmond, T.A., Gerhardt, D.J., Himmelbach, A., Clissold, L., Sampath, D., Ayling, S., Steuernagel, B., Pfeifer, M., D'Ascenzo, M., Akhunov, E.D., Hedley, P.E., Gonzales, A.M., Morrell, P.L., Kilian, B., Blattner, F.R., Scholz, U., Mayer, K.F., Flavell, A.J., Muehlbauer, G.J., Waugh, R., Jeddeloh, J.A., and Stein, N. (2013b). Barley whole exome capture: a tool for genomic research in the genus Hordeum and beyond. Plant J **76,** 494-505.

Matsumoto, T., Tanaka, T., Sakai, H., Amano, N., Kanamori, H., Kurita, K., Kikuta, A., Kamiya, K., Yamamoto, M., Ikawa, H., Fujii, N., Hori, K., Itoh, T., and Sato, K. (2011). Comprehensive sequence analysis of 24,783 barley full-length cDNAs derived from 12 clone libraries. Plant Physiol **156,** 20-28.

Matthews, D.E., Carollo, V.L., Lazo, G.R., and Anderson, O.D. (2003). GrainGenes, the genome database for small-grain crops. Nucleic Acids Res **31,** 183-186.

Mayer, K.F., Taudien, S., Martis, M., Simkova, H., Suchankova, P., Gundlach, H., Wicker, T., Petzold, A., Felder, M., Steuernagel, B., Scholz, U., Graner, A., Platzer, M., Dolezel, J., and Stein, N. (2009). Gene content and virtual gene order of barley chromosome 1H. Plant Physiol **151,** 496-505.

Mayer, K.F., Martis, M., Hedley, P.E., Simkova, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., Kubalakova, M., Suchankova, P., Murat, F., Felder, M., Nussbaumer, T., Graner, A., Salse, J., Endo, T., Sakai, H., Tanaka, T., Itoh, T., Sato, K., Platzer, M., Matsumoto, T., Scholz, U., Dolezel, J., Waugh, R., and Stein, N. (2011). Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell **23,** 1249-1263.

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. Plant Physiol **123,** 439-442.

McClintock, B. (1951). Chromosome organization and genic expression. Cold Spring Harbor symposia on quantitative biology **16,** 13-47.

Meins, F., Jr., Si-Ammour, A., and Blevins, T. (2005). RNA silencing systems and their relevance to plant development. Annual review of cell and developmental biology **21,** 297-318.

Miura, E., Kato, Y., Matsushima, R., Albrecht, V., Laalami, S., and Sakamoto, W. (2007). The balance between protein synthesis and degradation in chloroplasts determines leaf variegation in Arabidopsis yellow variegated mutants. Plant Cell **19,** 1313-1328.

Monaco, M.K., Stein, J., Naithani, S., Wei, S., Dharmawardhana, P., Kumari, S., Amarasinghe, V., Youens-Clark, K., Thomason, J., Preece, J., Pasternak, S., Olson, A., Jiao, Y., Lu, Z., Bolser, D., Kerhornou, A., Staines, D., Walts, B., Wu, G., D'Eustachio, P., Haw, R., Croft, D., Kersey, P.J., Stein, L., Jaiswal, P., and Ware, D. (2014). Gramene 2013: comparative plant genomics resources. Nucleic Acids Res **42,** D1193-1199.

Morell, M.K., Kosar-Hashemi, B., Cmiel, M., Samuel, M.S., Chandler, P., Rahman, S., Buleon, A., Batey, I.L., and Li, Z. (2003). Barley sex6 mutants lack starch synthase IIa activity and contain a starch with novel properties. Plant J **34,** 173-185.

Mount, D.W. (2007). Using the Basic Local Alignment Search Tool (BLAST). CSH protocols **2007,** pdb top17.

Muhlhausen, S., and Kollmar, M. (2013). Whole genome duplication events in plant evolution reconstructed and predicted using myosin motor proteins. BMC Evol Biol **13,** 202.

Muller, K.J., Romano, N., Gerstner, O., Garcia-Maroto, F., Pozzi, C., Salamini, F., and Rohde, W. (1995). The barley Hooded mutation caused by a duplication in a homeobox gene intron. Nature **374,** 727-730.

Naested, H., Holm, A., Jenkins, T., Nielsen, H.B., Harris, C.A., Beale, M.H., Andersen, M., Mant, A., Scheller, H., Camara, B., Mattsson, O., and Mundy, J. (2004). Arabidopsis VARIEGATED 3 encodes a chloroplast-targeted, zinc-finger protein required for chloroplast and palisade cell development. J Cell Sci **117,** 4807-4818.

Nagy, B., Majer, P., Mihály, R., Dudits, D., and Horváth, G.V. (2011). Transient and transgenic approaches for functional testing of candidate genes in barley. Acta Biologica Szegediensis **55,** 129-133.

Nair, S.K., Wang, N., Turuspekov, Y., Pourkheirandish, M., Sinsuwongwat, S., Chen, G., Sameri, M., Tagiri, A., Honda, I., Watanabe, Y., Kanamori, H., Wicker, T., Stein, N., Nagamura, Y., Matsumoto, T., and Komatsuda, T. (2010). Cleistogamous flowering in barley arises from the suppression of microRNA-guided HvAP2 mRNA cleavage. Proc Natl Acad Sci U S A **107,** 490-495.

Nelson, B.K., Cai, X., and Nebenfuhr, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant J **51,** 1126-1136.

Neuhaus, H.E., and Emes, M.J. (2000). Nonphotosynthetic Metabolism in Plastids. Annu Rev Plant Physiol Plant Mol Biol **51,** 111-140.

Newton, K.J., and Coe, E.H. (1986). Mitochondrial DNA changes in abnormal growth (nonchromosomal stripe) mutants of maize. Proc Natl Acad Sci U S A **83,** 7363-7366.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein engineering **10,** 1-6.

Nussbaumer, T., Martis, M.M., Roessner, S.K., Pfeifer, M., Bader, K.C., Sharma, S., Gundlach, H., and Spannagl, M. (2013). MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Res **41,** D1144-1151.

O'Sullivan, H. (2007). GrainGenes. Methods Mol Biol **406,** 301-314.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-i., Inokuchi, H., and Ozeki, H. (1986). Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature **322,** 572-574.

Okada, K., Saito, T., Nakagawa, T., Kawamukai, M., and Kamiya, Y. (2000). Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in Arabidopsis. Plant Physiol **122,** 1045-1056.

Oleykowski, C.A., Bronson Mullins, C.R., Godwin, A.K., and Yeung, A.T. (1998). Mutation detection using a novel plant endonuclease. Nucleic Acids Res **26,** 4597-4602.

Osborn, M.J., Panoskaltsis-Mortari, A., McElmurry, R.T., Bell, S.K., Vignali, D.A., Ryan, M.D., Wilber, A.C., McIvor, R.S., Tolar, J., and Blazar, B.R. (2005). A picornaviral 2A-like sequence-based tricistronic vector allowing for high-level therapeutic gene expression coupled to a dual-reporter system. Molecular therapy : the journal of the American Society of Gene Therapy **12,** 569-574.

Palmer, J.D. (1985). Comparative organization of chloroplast genomes. Annual review of genetics **19,** 325-354.

Park, S., and Rodermel, S.R. (2004). Mutations in ClpC2/Hsp100 suppress the requirement for FtsH in thylakoid membrane biogenesis. Proc Natl Acad Sci U S A **101,** 12765-12770.

Paterson, A.H., Bowers, J.E., and Chapman, B.A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A **101,** 9903-9908.

Pesaresi, P., Schneider, A., Kleine, T., and Leister, D. (2007). Interorganellar communication. Curr Opin Plant Biol **10,** 600-606.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8,** 785-786.

Petsalaki, E.I., Bagos, P.G., Litou, Z.I., and Hamodrakas, S.J. (2006). PredSL: a tool for the N-terminal sequence-based prediction of protein subcellular localization. Genomics, proteomics & bioinformatics **4,** 48-55.

Poland, J.A., and Rife, T.W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. Plant Genome-Us **5,** 92-102.

Poland, J.A., Brown, P.J., Sorrells, M.E., and Jannink, J.L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One **7,** e32253.

Potokina, E., Druka, A., Luo, Z., Wise, R., Waugh, R., and Kearsey, M. (2008). Gene expression quantitative trait locus analysis of 16 000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. Plant J **53,** 90-101.

Puchta, H., and Fauser, F. (2013). Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant J.

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell **80,** 847-857.

Pyke, K.A., and Leech, R.M. (1992). Chloroplast Division and Expansion Is Radically Altered by Nuclear Mutations in Arabidopsis thaliana. Plant Physiol **99,** 1005-1008.

Ramirez-Gonzalez, R.H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., Jack, P., Caccamo, M., and Uauy, C. (2014). RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant Biotechnol J.

Raven, J.A., and Allen, J.F. (2003). Genomics and chloroplast evolution: what did cyanobacteria do for plants? Genome biology **4,** 209.

Reichenbächer, D., Börner, T., and Richter, J. (1978). Untersuchungen am Fraktion-I-Protein der Gerste mit Hilfe quantitativer Immunelektrophoresen. Biochem. Physiol. Pflanzen **172,** 53-60.

Robert, L.S., Robson, F., Sharpe, A., Lydiate, D., and Coupland, G. (1998). Conserved structure and function of the Arabidopsis flowering time gene CONSTANS in Brassica napus. Plant Mol Biol **37,** 763-772.

Rodermel, S. (2002). Arabidopsis variegation mutants. Arabidopsis Book **1,** e0079.

Rostoks, N., Schmierer, D., Mudie, S., Drader, T., Brueggeman, R., Caldwell, D.G., Waugh, R., and Kleinhofs, A. (2006). Barley necrotic locus nec1 encodes the cyclic nucleotide-gated ion channel 4 homologous to the Arabidopsis HLM1. Mol Genet Genomics **275,** 159-168.

Rostoks, N., Mudie, S., Cardle, L., Russell, J., Ramsay, L., Booth, A., Svensson, J.T., Wanamaker, S.I., Walia, H., Rodriguez, E.M., Hedley, P.E., Liu, H., Morris, J., Close, T.J., Marshall, D.F., and Waugh, R. (2005). Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Mol Genet Genomics **274,** 515-527.

Ruppel, N.J., Kropp, K.N., Davis, P.A., Martin, A.E., Luesse, D.R., and Hangarter, R.P. (2013). Mutations in GERANYLGERANYL DIPHOSPHATE SYNTHASE 1 affect chloroplast development in Arabidopsis thaliana (Brassicaceae). Am J Bot **100,** 2074-2084.

Sager, R., and Ishida, M.R. (1963). Chloroplast DNA in Chlamydomonas. Proc Natl Acad Sci U S A **50,** 725-730.

Sakai, A., Takano, H., and Kuroiwa, T. (2004). Organelle nuclei in higher plants: structure, composition, function, and evolution. International review of cytology **238,** 59-118.

Sakamoto, W., Kondo, H., Murata, M., and Motoyoshi, F. (1996). Altered mitochondrial gene expression in a maternal distorted leaf mutant of Arabidopsis induced by chloroplast mutator. Plant Cell **8,** 1377-1390.

Sakamoto, W., Tamura, T., Hanba-Tomita, Y., Murata, M., and Sodmergen. (2002). The VAR1 locus of Arabidopsis encodes a chloroplastic FtsH and is

responsible for leaf variegation in the mutant alleles. Genes to cells : devoted to molecular & cellular mechanisms **7,** 769-780.

Saleh, A., Alvarez-Venegas, R., and Avramova, Z. (2008). An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. Nature protocols **3,** 1018-1025.

SAMBROOK, J., FRITSCH, E.F., and MANIATIS, T. (1989). Molecular Cloning: A Laboratory Manual, vol. I. 2nd edition. Cold Spring Harbor Laboratory Press. ISBN 0-87969-309-6.

Saski, C., Lee, S.B., Fjellheim, S., Guda, C., Jansen, R.K., Luo, H., Tomkins, J., Rognli, O.A., Daniell, H., and Clarke, J.L. (2007). Complete chloroplast genome sequences of Hordeum vulgare, Sorghum bicolor and Agrostis stolonifera, and comparative analyses with other grass genomes. Theor Appl Genet **115,** 571-590.

Sato, K., Nankaku, N., and Takeda, K. (2009a). A high-density transcript linkage map of barley derived from a single population. Heredity (Edinb) **103,** 110-117.

Sato, K., Shin, I.T., Seki, M., Shinozaki, K., Yoshida, H., Takeda, K., Yamazaki, Y., Conte, M., and Kohara, Y. (2009b). Development of 5006 full-length CDNAs in barley: a tool for accessing cereal genomics resources. DNA Res **16,** 81-89.

Schneeberger, K. (2014). Using next-generation sequencing to isolate mutant genes from forward genetic screens. Nat Rev Genet **15,** 662-676.

Schulte, D., Ariyadasa, R., Shi, B., Fleury, D., Saski, C., Atkins, M., deJong, P., Wu, C.C., Graner, A., Langridge, P., and Stein, N. (2011). BAC library resources for map-based cloning and physical map construction in barley (Hordeum vulgare L.). BMC Genomics **12,** 247.

Schwacke, R., Fischer, K., Ketelsen, B., Krupinska, K., and Krause, K. (2007). Comparative survey of plastid and mitochondrial targeting properties of transcription factors in Arabidopsis and rice. Mol Genet Genomics **277,** 631-646.

Seibel, N.M., Eljouni, J., Nalaskowski, M.M., and Hampe, W. (2007). Nuclear localization of enhanced green fluorescent protein homomultimers. Analytical biochemistry **368,** 95-99.

Serino, G., and Maliga, P. (1998). RNA polymerase subunits encoded by the plastid rpo genes are not shared with the nucleus-encoded plastid enzyme. Plant Physiol **117,** 1165-1170.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermuller, J., Reinhardt, R., Stadler, P.F., and Vogel, J. (2010). The primary transcriptome of the major human pathogen Helicobacter pylori. Nature **464,** 250-255.

Shen, L., Gong, J., Caldo, R.A., Nettleton, D., Cook, D., Wise, R.P., and Dickerson, J.A. (2005). BarleyBase--an expression profiling database for plant genomics. Nucleic Acids Res **33,** D614-618.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J **5,** 2043-2049.

Shirasu, K., Lahaye, T., Tan, M.W., Zhou, F., Azevedo, C., and Schulze-Lefert, P. (1999). A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in C. elegans. Cell **99,** 355-366.

Simons, G., van der Lee, T., Diergaarde, P., van Daelen, R., Groenendijk, J., Frijters, A., Buschges, R., Hollricher, K., Topsch, S., Schulze-Lefert, P., Salamini, F., Zabeau, M., and Vos, P. (1997). AFLP-based fine mapping of the Mlo gene to a 30-kb DNA segment of the barley genome. Genomics **44,** 61-70.

Small, I., Wintz, H., Akashi, K., and Mireau, H. (1998). Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol Biol **38,** 265-277.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics **4,** 1581-1590.

Spannagl, M., Noubibou, O., Haase, D., Yang, L., Gundlach, H., Hindemitt, T., Klee, K., Haberer, G., Schoof, H., and Mayer, K.F. (2007). MIPSPlantsDB--plant database resource for integrative and comparative plant genome research. Nucleic Acids Res **35,** D834-840.

Sparks, C.A., and Jones, H.D. (2009). Biolistics transformation of wheat. Methods Mol Biol **478,** 71-92.

Stein, N., and Graner, A. (2005). Map-based gene isolation in cereal genomes. In: Gupta PK, Varshney RK (Eds) Cereal genomics. Springer, Netherlands,, pp 331–360.

Stein, N., Perovic, D., Kumlehn, J., Pellio, B., Stracke, S., Streng, S., Ordon, F., and Graner, A. (2005). The eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in Hordeum vulgare (L.). Plant J **42,** 912-922.

Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H., Wolf, M., Kota, R., Varshney, R.K., Perovic, D., Grosse, I., and Graner, A. (2007). A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theoretical and Applied Genetics **114,** 823-839.

Stern, D.B., Goldschmidt-Clermont, M., and Hanson, M.R. (2010). Chloroplast RNA metabolism. Annu Rev Plant Biol **61,** 125-155.

Strand, A., Asami, T., Alonso, J., Ecker, J.R., and Chory, J. (2003). Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. Nature **421,** 79-83.

Strayer, C., Oyama, T., Schultz, T.F., Raman, R., Somers, D.E., Mas, P., Panda, S., Kreps, J.A., and Kay, S.A. (2000). Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. Science **289,** 768-771.

Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature **410,** 1116-1120.

Sugiura, M. (1992). The chloroplast genome. Plant Mol Biol **19,** 149-168.

Sun, C.W., Huang, Y.C., and Chang, H.Y. (2009). CIA2 coordinately up-regulates protein import and synthesis in leaf chloroplasts. Plant Physiol **150,** 879-888.

Sun, C.W., Chen, L.J., Lin, L.C., and Li, H.M. (2001). Leaf-specific upregulation of chloroplast translocon genes by a CCT motif-containing protein, CIA 2. Plant Cell **13,** 2053-2061.

Sun, X., Feng, P., Xu, X., Guo, H., Ma, J., Chi, W., Lin, R., Lu, C., and Zhang, L. (2011). A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. Nature communications **2,** 477.

Sunderland, P.A., West, C.E., Waterworth, W.M., and Bray, C.M. (2006). An evolutionarily conserved translation initiation mechanism regulates nuclear or mitochondrial targeting of DNA ligase 1 in Arabidopsis thaliana. Plant J **47,** 356-367.

Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990). Stable transformation of plastids in higher plants. Proc Natl Acad Sci U S A **87,** 8526-8530.

Swiatecka-Hagenbruch, M., Liere, K., and Borner, T. (2007). High diversity of plastidial promoters in Arabidopsis thaliana. Mol Genet Genomics **277,** 725-734.

Taketa, S., Yuo, T., Tonooka, T., Tsumuraya, Y., Inagaki, Y., Haruyama, N., Larroque, O., and Jobling, S.A. (2012). Functional characterization of barley betaglucanless mutants demonstrates a unique role for CslF6 in (1,3;1,4)-beta-D-glucan biosynthesis. J Exp Bot **63,** 381-392.

Taketa, S., Amano, S., Tsujino, Y., Sato, T., Saisho, D., Kakeda, K., Nomura, M., Suzuki, T., Matsumoto, T., Sato, K., Kanamori, H., Kawasaki, S., and Takeda, K. (2008). Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. Proc Natl Acad Sci U S A **105,** 4062-4067.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30,** 2725-2729.

Thiel, T., Graner, A., Waugh, R., Grosse, I., Close, T.J., and Stein, N. (2009). Evidence and evolutionary analysis of ancient whole-genome duplication in barley predating the divergence from rice. BMC Evol Biol **9,** 209.

Toshoji, H., Katsumata, T., Takusagawa, M., Yusa, Y., and Sakai, A. (2012). Effects of chloroplast dysfunction on mitochondria: white sectors in variegated leaves have higher mitochondrial DNA levels and lower dark respiration rates than green sectors. Protoplasma **249,** 805-817.

Tsugane, K., Maekawa, M., Takagi, K., Takahara, H., Qian, Q., Eun, C.H., and Iida, S. (2006). An active DNA transposon nDart causing leaf variegation and mutable dwarfism and its related elements in rice. Plant J **45,** 46-57.

Turner, A., Beales, J., Faure, S., Dunford, R.P., and Laurie, D.A. (2005). The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science **310,** 1031-1034.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res **40,** e115.

Valouev, A., Johnson, D.S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., Myers, R.M., and Sidow, A. (2008). Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. Nat Methods **5,** 829-834.

Valverde, F. (2011). CONSTANS and the evolutionary origin of photoperiodic timing of flowering. J Exp Bot **62,** 2453-2463.

van Bel, A.J., Hibberd, J., Prufer, D., and Knoblauch, M. (2001). Novel approach in plastid transformation. Current opinion in biotechnology **12,** 144-149.

Van Ooijen, J.W. (2006). JoinMap ® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Netherlands.

Varshney, R.K., Marcel, T.C., Ramsay, L., Russell, J., Roder, M.S., Stein, N., Waugh, R., Langridge, P., Niks, R.E., and Graner, A. (2007). A high density barley microsatellite consensus map with 775 SSR loci. Theor Appl Genet **114,** 1091-1103.

Voorrips, R.E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. The Journal of heredity **93,** 77-78.

Voytas, D.F. (2013). Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol **64,** 327-350.

Wagner, R., and Pfannschmidt, T. (2006). Eukaryotic transcription factors in plastids--Bioinformatic assessment and implications for the evolution of gene expression machineries in plants. Gene **381,** 62-70.

Walbot, V., and Coe, E.H. (1979). Nuclear gene iojap conditions a programmed change to ribosome-less plastids in Zea mays. Proc Natl Acad Sci U S A **76,** 2760-2764.

Wendt, T., Holm, P.B., Starker, C.G., Christian, M., Voytas, D.F., Brinch-Pedersen, H., and Holme, I.B. (2013). TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. Plant Mol Biol **83,** 279-285.

Wenzl, P., Li, H., Carling, J., Zhou, M., Raman, H., Paul, E., Hearnden, P., Maier, C., Xia, L., Caig, V., Ovesna, J., Cakir, M., Poulsen, D., Wang, J., Raman, R., Smith, K.P., Muehlbauer, G.J., Chalmers, K.J., Kleinhofs, A., Huttner, E., and Kilian, A. (2006). A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics **7,** 206.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Muller, K.F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol **76,** 273-297.

Wing, R.A., Zhang, H.B., and Tanksley, S.D. (1994). Map-based cloning in crop plants. Tomato as a model system: I. Genetic and physical mapping of jointless. Mol Gen Genet **242,** 681-688.

Wise, R.P., Caldo, R.A., Hong, L., Shen, L., Cannon, E., and Dickerson, J.A. (2007). BarleyBase/PLEXdb. Methods Mol Biol **406,** 347-363.

Wu, D., Wright, D.A., Wetzel, C., Voytas, D.F., and Rodermel, S. (1999). The IMMUTANS variegation locus of Arabidopsis defines a mitochondrial alternative oxidase homolog that functions during early chloroplast biogenesis. Plant Cell **11,** 43-55.

Xiong, J.Y., Lai, C.X., Qu, Z., Yang, X.Y., Qin, X.H., and Liu, G.Q. (2009). Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. Plant Mol Biol **71,** 437-449.

Yagi, Y., and Shiina, T. (2014). Recent advances in the study of chloroplast gene expression and its evolution. Front Plant Sci **5,** 61.

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., Sanchez, A., Valarik, M., Yasuda, S., and Dubcovsky, J. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proc Natl Acad Sci U S A **103,** 19581-19586.

Yang, P., Lupken, T., Habekuss, A., Hensel, G., Steuernagel, B., Kilian, B., Ariyadasa, R., Himmelbach, A., Kumlehn, J., Scholz, U., Ordon, F., and

Stein, N. (2014). PROTEIN DISULFIDE ISOMERASE LIKE 5-1 is a susceptibility factor to plant viruses. Proc Natl Acad Sci U S A **111,** 2104-2109.

Yaronskaya, E., Ziemann, V., Walter, G., Averina, N., Borner, T., and Grimm, B. (2003). Metabolic control of the tetrapyrrole biosynthetic pathway for porphyrin distribution in the barley mutant albostrians. Plant J **35,** 512-522.

Yoo, H.H., Kwon, C., Lee, M.M., and Chung, I.K. (2007). Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. Plant J **49,** 442-451.

Youens-Clark, K., Buckler, E., Casstevens, T., Chen, C., Declerck, G., Derwent, P., Dharmawardhana, P., Jaiswal, P., Kersey, P., Karthikeyan, A.S., Lu, J., McCouch, S.R., Ren, L., Spooner, W., Stein, J.C., Thomason, J., Wei, S., and Ware, D. (2011). Gramene database in 2010: updates and extensions. Nucleic Acids Res **39,** D1085-1094.

Yu, F., Liu, X., Alsheikh, M., Park, S., and Rodermel, S. (2008). Mutations in SUPPRESSOR OF VARIEGATION1, a factor required for normal chloroplast translation, suppress var2-mediated leaf variegation in Arabidopsis. Plant Cell **20,** 1786-1804.

Yuo, T., Yamashita, Y., Kanamori, H., Matsumoto, T., Lundqvist, U., Sato, K., Ichii, M., Jobling, S.A., and Taketa, S. (2012). A SHORT INTERNODES (SHI) family transcription factor gene regulates awn elongation and pistil morphology in barley. J Exp Bot **63,** 5223-5232.

Zakhrabekova, S., Gough, S.P., Braumann, I., Muller, A.H., Lundqvist, J., Ahmann, K., Dockter, C., Matyszczak, I., Kurowska, M., Druka, A., Waugh, R., Graner, A., Stein, N., Steuernagel, B., Lundqvist, U., and Hansson, M. (2012). Induced mutations in circadian clock regulator Mat-a facilitated short-season adaptation and range extension in cultivated barley. Proc Natl Acad Sci U S A **109,** 4326-4331.

Zhang, C.Y., Wang, Z., Bruce, H.L., Janz, J., Goddard, E., Moore, S., and Plastow, G.S. (2014). Associations between single nucleotide polymorphisms in 33 candidate genes and meat quality traits in commercial pigs. Animal genetics.

Zhang, L., Lavery, L., Gill, U., Gill, K., Steffenson, B., Yan, G., Chen, X., and Kleinhofs, A. (2009a). A cation/proton-exchanging protein is a candidate for the barley NecS1 gene controlling necrosis and enhanced defense response to stem rust. Theor Appl Genet **118,** 385-397.

Zhang, L., Wei, Q., Wu, W., Cheng, Y., Hu, G., Hu, F., Sun, Y., Zhu, Y., Sakamoto, W., and Huang, J. (2009b). Activation of the heterotrimeric G protein alpha-subunit GPA1 suppresses the ftsh-mediated inhibition of chloroplast development in Arabidopsis. Plant J **58,** 1041-1053.

Zhelyazkova, P., Sharma, C.M., Forstner, K.U., Liere, K., Vogel, J., and Borner, T. (2012). The primary transcriptome of barley chloroplasts: numerous

noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. Plant Cell **24,** 123-136.

7 Appendix

Appendix Table 1: Structure of the mapping population.

Population ID	Total # of individuals	# of analyzed individuals	# of Recombinants
	91 mapping population		
MM4205-72_1	22	22	n.d
MM4205-130_1	20	20	n.d
MM4205-230_1	21	21	n.d
MM4205-231_1	21	21	n.d
MM4205-237_1	7	7	n.d
	960 mapping population[A]		
MM4205-72_1	192	192	23
MM4205-130_2	192	192	36
MM4205-230_13	192	192	27
MM4205-231_1	192	192	23
MM4205-237_3	192	192	33
	1,920 mapping population[B]		
MM4205-72_2	192	47	0
MM4205-72_4	192	41	0
MM4205-130_5	192	43	0
MM4205-130_7	192	50	0
MM4205-230_7	192	33	0
MM4205-230_9	192	33	1
MM4205-231_2	192	22	0
MM4205-231_4	192	37	0
MM4205-237_11	192	48	0
MM4205-237_12	192	30	0

[A]- The recombinants were selected using flanking markers CAPS_2536 and CAPS_2560.

[B]- The recombinants were selected using flanking markers Zip_2661 and Zip_2680.

n.d- Not determined.

Appendix Table 2: F_3 phenotyping to determine the *albostrians* genotype of original F_2 plants.

91 F_2 mapping population						
Plant family ID	Repetition	Green	Variegated & *albino*	Σ 1&2		Genotype score
				Green	Variegated & *albino*	
MM4205-72_1_3	1	16	0	31	0	A
	2	15	0			
MM4205-72_1_7	1	0	14	0	29	B
	2	0	15			
MM4205-72_1_14	1	10	3	21	6	H
	2	11	3			
MM4205-72_1_20	1	3	12	6	25	B
	2	3	13			
MM4205-72_1_26	1	12	4	27	5	H
	2	15	1			
MM4205-72_1_31	1	5	10	7	24	B
	2	2	14			
MM4205-72_1_33	1	0	16	0	31	B
	2	0	15			
MM4205-72_1_36	1	1	15	5	27	B
	2	4	12			
MM4205-72_1_37	1	12	4	24	7	H
	2	12	3			
MM4205-72_1_45	1	1	15	2	30	B
	2	1	15			
MM4205-72_1_52	1	4	12	4	28	B
	2	0	16			
MM4205-72_1_70	1	0	16	0	32	B
	2	0	16			
MM4205-72_1_91	1	13	0	26	0	A
	2	13	0			
MM4205-72_1_114	1	15	0	31	0	A
	2	16	0			
MM4205-72_1_118	1	6	9	12	19	B
	2	6	10			
MM4205-72_1_121	1	16	0	32	0	A
	2	16	0			
MM4205-72_1_150	1	14	2	24	8	H
	2	10	6			
MM4205-72_1_161	1	8	6	22	7	H
	2	14	1			
MM4205-72_1_180	1	11	3	20	9	H
	2	9	6			
MM4205-72_1_185	1	12	3	20	5	H
	2	8	2			
MM4205-72_1_186	1	16	0	31	0	A
	2	15	0			
MM4205-72_1_188	1	13	3	26	5	H
	2	13	2			
MM4205-72_1_189	1	12	3	22	7	H

	2	10	4			
MM4205-130_2_1	1	12	4	26	6	H
	2	14	2			
MM4205-130_2_2	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_4	1	13	3	25	7	H
	2	12	4			
MM4205-130_2_9	1	11	4	23	7	H
	2	12	3			
MM4205-130_2_13	1	15	0	30	1	H
	2	15	1			
MM4205-130_2_15	1	1	15	2	30	B
	2	1	15			
MM4205-130_2_20	1	11	1	22	3	H
	2	11	2			
MM4205-130_2_26	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_29	1	15	0	31	0	A
	2	16	0			
MM4205-130_2_31	1	12	4	26	6	H
	2	14	2			
MM4205-130_2_33	1	0	16	0	32	B
	2	0	16			
MM4205-130_2_35	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_36	1	6	4	16	10	H
	2	10	6			
MM4205-130_2_37	1	10	6	18	13	H
	2	8	7			
MM4205-130_2_38	1	2	14	4	28	B
	2	2	14			
MM4205-130_2_39	1	16	0	31	0	A
	2	15	0			
MM4205-130_2_61	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_64	1	8	4	21	7	H
	2	13	3			
MM4205-130_2_81	1	0	13	4	25	B
	2	4	12			
MM4205-130_2_86	1	1	15	1	31	B
	2	0	16			
MM4205-130_2_90	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_112	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_115	1	14	1	28	3	H
	2	14	2			
MM4205-130_2_127	1	12	4	17	15	H
	2	5	11			

MM4205-130_2_136	1	15	0	31	0	A
	2	16	0			
MM4205-130_2_143	1	14	1	25	6	H
	2	11	5			
MM4205-130_2_144	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_146	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_149	1	1	13	1	28	B
	2	0	15			
MM4205-130_2_152	1	12	4	23	9	H
	2	11	5			
MM4205-130_2_164	1	13	3	26	6	H
	2	13	3			
MM4205-130_2_165	1	4	12	8	24	B
	2	4	12			
MM4205-130_2_175	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_182	1	4	11	9	22	B
	2	5	11			
MM4205-130_2_183	1	16	0	32	0	A
	2	16	0			
MM4205-130_2_188	1	14	0	30	0	A
	2	16	0			
MM4205-230_13_3	1	14	2	26	3	H
	2	12	1			
MM4205-230_13_4	1	16	0	32	0	A
	2	16	0			
MM4205-230_13_18	1	15	1	31	1	H
	2	16	0			
MM4205-230_13_29	1	14	2	28	4	H
	2	14	2			
MM4205-230_13_40	1	16	0	31	0	A
	2	15	0			
MM4205-230_13_42	1	10	4	21	9	H
	2	11	5			
MM4205-230_13_55	1	16	0	31	0	A
	2	15	0			
MM4205-230_13_59	1	0	16	1	31	B
	2	1	15			
MM4205-230_13_82	1	11	5	23	9	H
	2	12	4			
MM4205-230_13_88	1	11	4	20	9	H
	2	9	5			
MM4205-230_13_92	1	14	2	27	5	H
	2	13	3			
MM4205-230_13_108	1	9	6	20	11	H
	2	11	5			
MM4205-230_13_122	1	3	13	5	27	B

	2	2	14			
MM4205-230_13_126	1	15	1	31	1	H
	2	16	0			
MM4205-230_13_139	1	12	4	21	8	H
	2	9	4			
MM4205-230_13_144	1	16	0	32	0	A
	2	16	0			
MM4205-230_13_149	1	12	3	25	6	H
	2	13	3			
MM4205-230_13_162	1	0	16	0	32	B
	2	0	16			
MM4205-230_13_164	1	0	16	1	30	B
	2	1	14			
MM4205-230_13_166	1	6	10	14	18	B
	2	8	8			
MM4205-230_13_168	1	1	15	2	30	B
	2	1	15			
MM4205-230_13_171	1	12	4	26	6	H
	2	14	2			
MM4205-230_13_174	1	0	15	0	30	B
	2	0	15			
MM4205-230_13_175	1	10	6	20	10	H
	2	10	4			
MM4205-230_13_176	1	1	13	2	27	B
	2	1	14			
MM4205-230_13_183	1	13	3	26	3	H
	2	13	0			
MM4205-230_13_190	1	8	8	14	15	B
	2	6	7			
MM4205-231_1_12	1	14	0	28	0	A
	2	14	0			
MM4205-231_1_17	1	13	2	25	6	H
	2	12	4			
MM4205-231_1_18	1	16	0	31	0	A
	2	15	0			
MM4205-231_1_21	1	9	7	18	13	H
	2	9	6			
MM4205-231_1_28	1	0	16	1	31	B
	2	1	15			
MM4205-231_1_29	1	0	15	0	31	B
	2	0	16			
MM4205-231_1_31	1	16	0	32	0	A
	2	16	0			
MM4205-231_1_61	1	15	0	30	0	A
	2	15	0			
MM4205-231_1_77	1	8	7	20	11	H
	2	12	4			
MM4205-231_1_92	1	0	13	1	27	B
	2	1	14			

MM4205-231_1_103	1	15	0	31	0	A
	2	16	0			
MM4205-231_1_105	1	0	16	0	21	B
	2	0	5			
MM4205-231_1_107	1	16	0	31	1	H
	2	15	1			
MM4205-231_1_110	1	16	0	32	0	A
	2	16	0			
MM4205-231_1_118	1	13	3	25	7	H
	2	12	4			
MM4205-231_1_146	1	13	2	25	6	H
	2	12	4			
MM4205-231_1_154	1	16	0	28	0	A
	2	12	0			
MM4205-231_1_170	1	13	3	26	4	H
	2	13	1			
MM4205-231_1_177	1	15	1	27	2	H
	2	12	1			
MM4205-231_1_179	1	16	0	32	0	A
	2	16	0			
MM4205-231_1_186	1	14	0	29	1	H
	2	15	1			
MM4205-231_1_192	1	0	16	0	28	B
	2	0	12			
MM4205-237_3_2	1	13	3	25	6	H
	2	12	3			
MM4205-237_3_3	1	14	2	25	4	H
	2	11	2			
MM4205-237_3_4	1	15	0	30	0	A
	2	15	0			
MM4205-237_3_6	1	2	14	5	26	B
	2	3	12			
MM4205-237_3_10	1	6	10	10	22	B
	2	4	12			
MM4205-237_3_22	1	13	3	25	7	H
	2	12	4			
MM4205-237_3_25	1	14	2	24	8	H
	2	10	6			
MM4205-237_3_34	1	15	0	22	0	A
	2	7	0			
MM4205-237_3_46	1	16	0	29	0	A
	2	13	0			
MM4205-237_3_72	1	6	10	10	22	B
	2	4	12			
MM4205-237_3_73	1	13	3	26	6	H
	2	13	3			
MM4205-237_3_75	1	12	4	22	9	H
	2	10	5			
MM4205-237_3_94	1	16	0	29	0	A

	2	13	0			
MM4205-237_3_96	1	11	5	24	7	H
	2	13	2			
MM4205-237_3_109	1	11	5	23	9	H
	2	12	4			
MM4205-237_3_110	1	4	10	7	17	B
	2	3	7			
MM4205-237_3_112	1	11	5	25	7	H
	2	14	2			
MM4205-237_3_122	1	16	0	30	0	A
	2	14	0			
MM4205-237_3_123	1	0	16	0	29	B
	2	0	13			
MM4205-237_3_134	1	14	2	25	6	H
	2	11	4			
MM4205-237_3_135	1	11	5	22	7	H
	2	11	2			
MM4205-237_3_139	1	6	1	6	1	H
	2	0	0			
MM4205-237_3_140	1	16	0	32	0	A
	2	16	0			
MM4205-237_3_145	1	11	5	21	10	H
	2	10	5			
MM4205-237_3_149	1	15	1	22	5	H
	2	7	4			
MM4205-237_3_151	1	11	5	23	8	H
	2	12	3			
MM4205-237_3_159	1	1	14	1	24	B
	2	0	10			
MM4205-237_3_178	1	1	15	4	26	B
	2	3	11			
MM4205-237_3_179	1	13	0	27	0	A
	2	14	0			
MM4205-237_3_180	1	13	0	25	0	A
	2	12	0			
MM4205-237_3_183	1	0	16	2	28	B
	2	2	12			
MM4205-237_3_189	1	1	15	2	25	B
	2	1	10			
MM4205-237_3_133	1	1	14	3	19	B
	2	2	5			
MM4205-230_13_142	1	2	14	5	27	B
	2	3	13			
MM4205-230_13_148	1	8	5	17	9	H
	2	9	4			
MM4205-231_1_172	1	12	1	24	3	H
	2	12	2			
MM4205-231_1_160	1	9	6	19	9	H
	2	10	3			

MM4205-237_3_54	1	3	12	4	27	B
	2	1	15			
MM4205-231_1_102	1	0	14	0	27	B
	2	0	13			
960 mapping population						
Plant family ID	Repetition	Green	Variegated & *albino*	Σ 1&2		Genotype score
				Green	Variegated & *albino*	
MM4205-72_1_1	1	15	0	30	0	A
	2	15	0			
MM4205-72_1_2	1	14	0	27	0	A
	2	13	0			
MM4205-72_1_3	1	13	2	25	5	H
	2	12	3			
MM4205-72_1_4	1	4	11	8	21	B
	2	4	10			
MM4205-72_1_5	1	12	3	24	6	H
	2	12	3			
MM4205-72_1_6	1	15	0	29	0	A
	2	14	0			
MM4205-72_1_7	1	0	15	0	30	B
	2	0	15			
MM4205-72_1_8	1	11	4	22	8	H
	2	11	4			
MM4205-72_1_9	1	3	12	7	23	B
	2	4	11			
MM4205-72_1_10	1	13	2	27	3	H
	2	14	1			
MM4205-72_1_11	1	15	0	30	0	A
	2	15	0			
MM4205-72_1_12	1	15	0	30	0	A
	2	15	0			
MM4205-72_1_13	1	13	1	25	4	H
	2	12	3			
MM4205-72_1_14	1	15	0	29	0	A
	2	14	0			
MM4205-72_1_15	1	3	11	3	26	B
	2	0	15			
MM4205-72_1_16	1	1	14	2	28	B
	2	1	14			
MM4205-72_1_17	1	11	4	24	6	H
	2	13	2			
MM4205-72_1_18	1	11	4	26	4	H
	2	15	0			
MM4205-72_1_19	1	0	15	0	30	B
	2	0	15			
MM4205-72_1_20	1	13	2	22	7	H
	2	9	5			
MM4205-72_1_21	1	0	15	0	30	B
	2	0	15			

MM4205-130_1_2	1	13	2	23	7	H
	2	10	5			
MM4205-130_1_3	1	15	0	29	0	A
	2	14	0			
MM4205-130_1_4	1	13	2	23	6	H
	2	10	4			
MM4205-130_1_5	1	13	0	27	0	A
	2	14	0			
MM4205-130_1_6	1	11	4	17	9	H
	2	6	5			
MM4205-130_1_7	1	10	5	23	7	H
	2	13	2			
MM4205-130_1_8	1	13	2	25	5	H
	2	12	3			
MM4205-130_1_9	1	15	0	30	0	A
	2	15	0			
MM4205-130_1_10	1	15	0	29	0	A
	2	14	0			
MM4205-130_1_11	1	11	4	20	10	H
	2	9	6			
MM4205-130_1_13	1	15	0	30	0	A
	2	15	0			
MM4205-130_1_14	1	11	4	22	8	H
	2	11	4			
MM4205-130_1_15	1	14	1	24	6	H
	2	10	5			
MM4205-130_1_16	1	11	4	23	7	H
	2	12	3			
MM4205-130_1_17	1	15	0	30	0	A
	2	15	0			
MM4205-130_1_18	1	15	0	30	0	A
	2	15	0			
MM4205-130_1_19	1	10	3	19	9	H
	2	9	6			
MM4205-130_1_20	1	12	3	22	8	H
	2	10	5			
MM4205-130_1_22	1	14	1	27	3	H
	2	13	2			
MM4205-230_1_1	1	15	0	30	0	A
	2	15	0			
MM4205-230_1_2	1	15	0	30	0	A
	2	15	0			
MM4205-230_1_3	1	15	0	30	0	A
	2	15	0			
MM4205-230_1_4	1	14	0	29	0	A
	2	15	0			
MM4205-230_1_5	1	0	15	0	30	B
	2	0	15			
MM4205-230_1_6	1	10	4	21	8	H

	2	11	4			
MM4205-230_1_7	1	9	4	19	9	H
	2	10	5			
MM4205-230_1_8	1	9	4	18	10	H
	2	9	6			
MM4205-230_1_9	1	15	0	30	0	A
	2	15	0			
MM4205-230_1_10	1	2	13	7	23	B
	2	5	10			
MM4205-230_1_11	1	15	0	30	0	A
	2	15	0			
MM4205-230_1_12	1	12	3	23	7	H
	2	11	4			
MM4205-230_1_13	1	12	3	26	4	H
	2	14	1			
MM4205-230_1_14	1	13	2	26	4	H
	2	13	2			
MM4205-230_1_15	1	2	13	8	22	B
	2	6	9			
MM4205-230_1_16	1	14	1	25	5	H
	2	11	4			
MM4205-230_1_17	1	15	0	30	0	A
	2	15	0			
MM4205-230_1_18	1	11	4	21	9	H
	2	10	5			
MM4205-230_1_19	1	15	0	24	6	H
	2	9	6			
MM4205-230_1_20	1	0	13	0	27	B
	2	0	14			
MM4205-230_1_21	1	14	1	28	2	H
	2	14	1			
MM4205-231_1_1	1	15	0	30	0	A
	2	15	0			
MM4205-231_1_2	1	7	5	18	8	H
	2	11	3			
MM4205-231_1_3	1	0	15	0	30	B
	2	0	15			
MM4205-231_1_5	1	15	0	29	0	A
	2	14	0			
MM4205-231_1_6	1	10	5	25	5	H
	2	15	0			
MM4205-231_1_7	1	9	6	19	10	H
	2	10	4			
MM4205-231_1_8	1	15	0	30	0	A
	2	15	0			
MM4205-231_1_10	1	10	5	19	10	H
	2	9	5			
MM4205-231_1_11	1	3	12	8	21	B
	2	5	9			

MM4205-231_1_12	1	11	3	18	8	H
	2	7	5			
MM4205-231_1_13	1	15	0	29	0	A
	2	14	0			
MM4205-231_1_15	1	14	0	29	0	A
	2	15	0			
MM4205-231_1_16	1	8	7	22	8	H
	2	14	1			
MM4205-231_1_18	1	9	6	18	10	H
	2	9	4			
MM4205-231_1_20	1	12	3	21	8	H
	2	9	5			
MM4205-231_1_21	1	12	3	23	6	H
	2	11	3			
MM4205-231_1_22	1	12	3	22	6	H
	2	10	3			
MM4205-237_1_1	1	1	14	4	26	B
	2	3	12			
MM4205-237_1_2	1	13	1	25	2	H
	2	12	1			
MM4205-237_1_3	1	13	2	25	4	H
	2	12	2			
MM4205-237_1_4	1	14	1	26	4	H
	2	12	3			
MM4205-237_1_5	1	14	0	29	0	A
	2	15	0			
MM4205-237_1_6	1	3	9	10	16	B
	2	7	7			
MM4205-237_1_7	1	1	14	7	23	B
	2	6	9			
MM4205-237_1_1	1	1	14	4	26	B
	2	3	12			

Appendix Table 3: Summary of primers used in this study for marker development.

Order	Primer ID	Fragment Size (bp)	Forward Primer	Reverse Primer
1	CAPS_2471 (k08084)	1885	GACGGTGGCCAGATAACAGT	AACAGCAGGAAGGAGCTGAA
2	CAPS_2500 (k02741)	450	TGGAAGTGTTGACTCCCTCA	GAAGGTGGTTTCAGCCATGT
3	CAPS_2503 (k06406)	790	GTGAAGGCACATCACAAAGC	TGAGCAGTATGCAATCAGGG
4	CAPS_2504 (k07980)	273	CTGCACAGCAAAATTCATGG	TCTCATTCAAGGTGTGCGAG
5	CAPS_2505 (k02916)	1076	GAAATGAAAACTAGCTCCCAGC	TCAGCGAAACATGATGTGGT
6	CAPS_2506 (k02987)	1213	GATTGGCCTTATGGTAGCGA	GGAGCGTTCTTTTGACAGGA
7	CAPS_2507 (k02572)	352	CGCCAACAATTTCAGCTTCT	TTCGGCACAACTACCTCACA
8	CAPS_2509 (k08641)	500	ATGGTGACAACAACCGACAA	CAGCACTTTGAACAGGAGGG
9	CAPS_2510 (k04173)	402	CCAGGTACCAAATGGGAGAA	AAGAAGCAGAAGGCAAACGA
10	CAPS_2511 (k07126)	268	AGAATAGGGGGCAGGTGTTT	TGTGGTATGAGCACCTTGGA
11	CAPS_2512 (k06889)	359	GGCTTTCTTGCATACCTGGA	CGTCGCTGGTAAATCTGGAG
12	CAPS_2513 (k07372)	275	CTTTCCAAGGCATCAACCAT	AACGCCTACTACCGGAACCT
13	CAPS_2514 (k07182)	634	ATCATCAGGGTGAGATGCAA	GCAGCAAGTGATTTCGATGA
14	CAPS_2515 (k08155)	920	ACACGAGCCCTCTTGCTATC	TGAGAGCATACCAGGGGAAC
15	CAPS_2523 (k06681)	400	CCATGTTCCAAACTCCAGGT	TACCAATCCTTCAAGCCGTC
16	CAPS_2524 (k00595)	366	GCTCCCTACTAAACAGGCCC	GAGCCCTGGAGACCATGATA
17	CAPS_2527 (k08467)	528	TTCGCTTACATCAACTTGCG	GAGGCTGTGGTTAAGCTTGC
18	CAPS_2536 (k08949)	734	CATCACGTAAATCAAGTGCCA	TTATGCACCTCCAATGCGTA
19	CAPS_2537 (k05223)	420	CAATACCATTACCGTTCCGC	GCAGCACCTGCAGAAAGAG
20	CAPS_2541 (k04151)	388	AGCAGAATACAATGCCAGGG	CGAGCGTGCTTATTTCATCA
21	CAPS_2542 (k09173)	404	CAGCGTTCAGATGTCTTGCT	GCACTGAGTCCGTTTCCATT
22	CAPS_2546 (k07221)	447	AGCAACAGCAACATGAGCAG	TCAACTATGACCTCCCGACC
23	CAPS_2547 (k09929)	354	GATCACATCTGACGTCCTCG	AGAAAAGGAAGGCAAGGAGC
24	CAPS_2548 (k04467)	349	GCTATCAGTGCGCATTACCA	GATTCGGTGCGGACAACTAT
25	CAPS_2549 (k03809)	381	CATCCATGATCTCTGCCCTT	CTGCATGGGAGATGTAAGAGG
26	CAPS_2551 (k03358)	444	ATCATCGATCCAAGCACTCC	CAAGGCGGAGATCTGACAAT
27	CAPS_2552 (k03428)	419	TTACCGAATTACGTCGGCTC	CAGCCAGAAGAGACAGGGAC
28	CAPS_2553 (k07898)	376	GATTGCAGGATTACATGCGA	TTGCTGACCTTCTGTGATCG
29	CAPS_2557 (k03102)	821	TCCAGTGCTCTTCGATCCTT	AAGGAGTCAGCCACGAGAAA
30	CAPS_2558 (k00397)	468	ACTTCGACAAAGGACGCTGT	CTACTACCGCCAGAAGCCAC
31	CAPS_2560 (k09456)	673	ACGACGACGAAGGTACTGCT	TCTATTCCTCAGTGGCCTCG
32	CAPS_2561 (k02044)	751	AATTGAGCCATTGCTTCTGG	CGATGATGGGGAAAGCTAAA
33	CAPS_2562 (k06299)	372	ACATCCTTGCTTGCTTGCTT	CACCCAGTGGCTCTGGTACT
34	CAPS_2563 (k07450)	252	TACAGACACCCCCATCTTCC	GGCGAATGGTACAATCAGGA
35	CAPS_2564 (k06628)	268	CGGGACTCATTTACGAAGTAGC	GGAGGAAAATAGGCACGGAT
36	CAPS_2565 (k08205)	378	GGAAAAACAGAGAGCATGGC	GAAACTTTTTCCTCCCAGCC
37	CAPS_2568 (k07247)	1608	CGACGATGCATAATGCTCAA	GGCTCCAGCTAGGGTTTCTC
38	CAPS_2571 (k04885)	383	TAGCATGGGGAAAACCTGAC	ACTTGTGGGCAAAAGCAATC
39	CAPS_2574 (k03027)	529	ACATTGAGAGCATGTGCAGC	AAGATGCATGTTAGGGTCGG
40	CAPS_2575 (k04949)	752	TCAGCTCCATCAACACTTCG	AGCTGGTGATAGGGCATTTG
41	CAPS_2578 (k06499)	1300	GCCGAAATACGTCTCACACA	CACCTCTCCGACGAACTCTC
42	CAPS_2586 (k02011)	350	ATCAGAAGCCATGATGTCCC	TTGTTCTCCACACTGATGCC
43	1_1051	572	ACCTACCAGGACCCTTCGAC	GGCCTAATGTTCCCCATCTT

44	2_0808	799	CAGATGGTCACAAGCAAGGA	TTCAGCAGCATAACCAGCAC
45	3_0168	783	CCCGGCTAAGTTCTGTCAAG	GACTAGGGAAACCTGCGACA
46	3_1188	756	CATGGTCCTGATCCAGCTTT	TGTGGTTACGGTCATCCTGA
47	3_1395	513	GCAGATCATCAGTGGGACCT	GTGTGTTTGCGTCTGTTGCT
48	Zip_2582	686	GGCTGCAAAAAGGAGTCAAG	AGCAAGTACTTCCGGAGCAA
49	Zip_2586	854	GGCATGAACCAGATCGTCTT	AGCCACTGAGCAGGATAGGA
50	Zip_2587	793	CCCTCTACAGTCCGATCTGC	CAGATGTTCTGGAGCGCATA
51	Zip_2588	698	CGACCTTGTCCCACTTCATT	CGGCTAATAACGGCTCTACG
52	Zip_2594	702	CATGAAAGGGGGAGATCAGA	GCTCTAAAGCGAACCAAACG
53	Zip_2601	810	TCTTGGCCGACTCTTATTGG	AGCAGCATCTTCAGCTCCTC
54	Zip_2603	865	CCAGCAGGCATGAGTGACTA	AGGAACGTGGTCACTTGGAC
55	Zip_2609	667	CACGAGGTGCTTATGTGGAA	AGACGCCTTCTTCATCTCCA
56	Zip_2613	854	ACGGTTGTAGACGGGTTGAG	ATGGGATGTGTAGGCCATGT
57	Zip_2625	690	GCTAACCAGAGCGAGCAGAG	GTGCCATGTGAGAAATGCAG
58	Zip_2626	708	GTTGTGTGGCCTTTCTGGTT	GCTCTTCCCTTACGTTGCTG
59	Zip_2630	809	ATGAAAGTGGATGGCCTGAC	GTTCAGCTCCGTCTTTCGAC
60	Zip_2631	854	GATCTGATCTGCGGATGGTT	ACGATCGAAACCCAAGTCAC
61	Zip_2634	734	ATGAGGATGATCCGACAAGC	CTCCACCCTGAGCTCACTTC
62	Zip_2641	553	AGCAACAAGATCCGGGACAT	CTCGGTCTGTTACCGGTGTT
63	Zip_2643	775	CATGTTCACCCCTTCTGGAT	GCTCCTTCACCTTTGTGAGC
64	Zip_2644	811	GTGGGTTTCAACCACTTGCT	GCAGAGCATCATCAGGACAA
65	Zip_2655	790	GCTGTGCTTTTTCCCTTGAG	CTGCAGGCACGATAGCAATA
66	Zip_2656	797	AGCAAGACTCCTCCACCAGA	CAAGGACTGTGGTGTCATGG
67	Zip_2661	843	GGATCTCTTGCAGAGCCAAC	ATCGCCACAAGGTTATCAGC
68	Zip_2662	699	CAGCGATGCTTCGACTATCA	GATGTGATGCCCTGTTTTCC
69	Zip_2665	959	CAACTGGAGTGGTTGGATTG	CCTAGCCGGAAAGAAGCTC
70	Zip_2667_1	954	GCCATCTTCCTCCTCATCAC	ATGGTCTGCCACATGGTGTA
71	Zip_2667_2	932	CCATTGCCATGCTTCCAT	CGGTCCAATACAAGATGCTC
72	Zip_2667_3	946	CCTCATTCCTCAATGGTGCT	AGCATCTCCTGCCACAACAT
73	Zip_2667_4	692	CATGTCCCCACTGTATGATCC	GCCAAACAGGAAACACCA
74	Zip_2669_1	983	CACCCTACCCCTCCCAAGTA	GTCGAAGGGTCCTGGTAGGT
75	Zip_2669_2	947	GCACGCCTGAATGTACCATA	GGCAAATCTTCTCCTCGTCTC
76	Zip_2671_1	970	TTTGCCATCAGCTCCTCTCT	ACAGCTCTCGGACTTCTGGA
77	Zip_2671_2	798	ACGGGGTATAGCACATTCCA	CAGGCTAACTGCCACATCAA
78	Zip_2672	613	AACGGAAACGCTCTCTCTCA	CAACCACCGCTGCTACTACA
79	Zip_2680_1	850	GCACTCAAGAGCAACATCCT	AGAAGGGCGGATTGTCTTCT
80	Zip_2680_2	645	GTCGTTGCCGATGTTGAAG	ACGTCGTGCACTACCATGAG
81	Zip_2682	646	CAACTCCTTGGGGAGCAATA	TTGGCATGAAGAGAGTGCTG
82	Zip_2687_1	677	CGGCTGAATTAGCATTGTGC	GCAACCATGTGTCACCAAGA
83	Zip_2687_2	655	GGTGACACATGGTTGCAGAA	GGGCACTGATATTTCGTCAC
84	Zip_2687_3	947	CAGGAGCTCTGAATCCGTATG	CCTATTTACGACCGCGAGAG
85	Zip_2689_1	641	CCTACATCCCCCATGTCATC	GTGGTTGGTGGTGAAGATCA
86	Zip_2689_2	609	CTGTCCGGTGTCCTCAACTT	ATCAAGGAGCCGCTTGTATG
87	Zip_2690	677	TCAGGAACATGGAGGGAAAG	CATCGAACACACAGGAGCAC
88	Contig_220966	669	GCTGCCACTCTTCACAATGA	GCCGTGGAGGATCAAACTAA

89	Contig_1648021	894	ATGTATCAGCATCCCCAAGC	GGCTAGCCATGGAATGTGTT
90	Contig_2565874	611	CAAACCGTTATGTGCACTGG	TCCACGGTGATACGTTGAAG
91	Contig_2548245	627	GCAGAATCCCAGTCAGCTTC	GAATCCTCCAGCTCTTGCAC
92	Contig_1586971	661	CGCACACAGTTCGTCTCATT	GTCGGGACCAAGACAACACT
93	Contig_1596897	683	TGGAGACGTATCACGCAGTC	GTACCCCTCGCCCTAAACTC
94	Contig_204826	744	GGATGTGATACTGGGGATGG	ATCGTCTATGCACACCACCA
95	Contig_47245	717	ATGTCAACCCTTGAGCCAAC	ACAGACGCGGGTATGAAAAC
96	Contig_52867	900	GTCCATTGCTGCTTCTGACA	GAGAAGCAAAAGCGTGAACC
97	Contig_7973	887	TATCACCTCGCTGCAAACAG	GCTTTTCCAAGAGCAACTGG
98	Contig_75832	772	CTTGTGGTTGTCGAGCTTCA	GCCAGTTCAGCCAAAGAAAG
99	Contig_39427	776	CGAGCTTTATTTCCCCTTCC	CATGGTCACATCGACACACA
100	Contig_139498	733	GGTTTGCCTCTTCAGCAAAG	GCACAGAGCCCTGAACTAGG
101	Contig_1590097	756	CATGGTCCTGATCCAGCTTT	TGTGGTTACGGTCATCCTGA
102	Contig_157273	885	CCAACCCTGCCACTGTATCT	AGCTCCGCCAATGTCTCTAA
103	Contig_72015	849	ATGTTGCCCCCTCTCTCTTT	CTAGCCCTCCGGAGAAGAAC
104	Contig_65509	753	GCACACACATTTTCCACCAG	TGGGTTTCTCTCAGGATTGG
105	Contig_441985	810	ACCATGGTGAAGGAGGACTG	CTATGGGCGACAAACAGGAT
106	Contig_137471	624	ACGGTGAAGGGGAGAAAAGT	CACTAAAGCCATCGTCAGCA
107	Contig_41240	760	TAGAACGGGAGCACCCTATG	CGGTTCTGGTGCCTTGTTAT
108	Contig_167530	605	CCATGTATGCCTCGACTGTG	TGGCTTAGGGATACCACTCG
109	Contig_72830	885	CTCATGAAGGACGTTGCAGA	GCCCTTTACATGCACCTGTT
110	Contig_141063	880	ACCCCTGGATCCATTACTCC	ATACGTTTCGGCAGTCAACC
111	Contig_40788	898	TGCGAGATAGCTGAGCTGAA	ACATGCCAGACCATCGTGTA
112	Contig_37799_1	850	CTCAGTGCTCCCTTGGTCAG	CACATCCACATGGCAGAAGA
113	Contig_37799_2	983	ATAGTTCCCCGTTGACGTGT	AGGTACAGGAAGTGGCACTCA
114	Contig_37799_3	949	GGGTTGCCTGGATATGTGTG	GGATGTCGTTGGACATGAGC
115	Contig_37799_4	967	AGCTCAGATCAATGGGCATC	CATCGTGCTTCTCTTTGCAG
116	Contig_37799_5	897	GAGCATCAAGCTGGTCACAA	GACTTTGATTGCCGGTGTGT
117	Contig_37952_1	850	GAGTAATACCCGCCCACAAA	CGAGTCGTCCTTGAAGTTGG
118	Contig_37952_2	922	CCTCTCCAACACCAGCAACT	ACATGGGAGAATGCAGAAGC
119	Contig_37952_3	967	GCTTCTGCATTCTCCCATGT	GTTCAGGTGGTGGTTGCTCT
120	Contig_37952_4	986	GGCAGGAATATCGGCAACTA	ACATAGGAGAGCAGGCGTCA
121	Contig_37952_5	849	ATCGGCAGAGATGAAACTGG	TGCCATGTACGAGCATAGGT
122	Contig_40728_1	976	GCCGTGCTAACATTTTGAGG	CGAGTTGTCACTAGCACTCTGC
123	Contig_40728_2	961	GGACCAAAGAGCCTACATGG	CCAAGAGGATGCAACTTGTG
124	Contig_40728_3	1172	CACAAGTTGCATCCTCTTGG	TGTGCACTATGTTGGGGTTG
125	Contig_49785_1	982	GGCAAGACGAAGTTATGACG	AGGCTTGCTGGTGGCTAATA
126	Contig_49785_2	947	GATGGCTCGGACATCAAATC	GAGCCTTAGGAGCAGCAGAA
127	Contig_49785_3	1196	TCTGAGGATGATGCAGATGG	TCCTGCAAATCGGAAGTAGG
128	Contig_49785_4	990	CCGATATTGATGCCTGTTCC	CGGTCAAGTCTGGCGTATTT
129	Contig_49785_5	790	GTTTAAGCGGGCTGTCAGAG	GTCTCATGGAGCCCAATCTC
130	Contig_104939	718	TGAGACCTCTCTGCCATGTG	TGTGGCTGCTAACGAATCTG
131	Contig_137310_1	1160	ACGAGCAGACTGTCATTGGA	CCGTGGAAGTAGACGAGCAC
132	Contig_137310_2	1180	ACGCCGAATATAACGACCAC	TAACCGACGAACACGTAGCA
133	Contig_274716	819	CTGCAGGGAAGATGCTAAGG	CGCAGTTGATCTAGCCACAA

134	Contig_275079	820	TGATTTGCAGTGAGGCTGTC	GCCCACAGGTTTGGAAGTTA
135	Contig_303042_1	812	TCTTTTGCCATCAGCTCCTC	GGGTAGCATTTGGAATGTGC
136	Contig_303042_2	798	ACGGGGTATAGCACATTCCA	CAGGCTAACTGCCACATCAA
137	Contig_1558472_1	996	AGCCAAGGGCTCACTTCTTT	CGATTCCAGTGGCAATGTAG
138	Contig_1558472_2	1019	GATCGCTGATCATCCACTGA	CATGCATGGGTAGGTTTGTC
139	Contig_1558472_3	1158	GGGCATCGTGTCAAATAGGT	GCTTACCTGCAGACGAATCC
140	Contig_1558472_4	803	GATCACCGTGAGCAGGAAAT	TAGGGCGAAACACTGAAACC
141	Contig_1561286_1	1190	TTAGGACCACCCTGGTTTCA	GCAGCATCCATGTCATTACC
142	Contig_1561286_2	1293	CGTGTCTGTGAAATGGTTGG	AGCGTTGTACTCCGGCTTTA
143	Contig_1563733_1	940	GTCCTCCTCACCTTTCACCA	GGAGACGTTGGAGTTGCCTA
144	Contig_1563733_2	916	GCCTAATCAATAACGCGGTGTC	TCGAAGAGGGAAGCTTGAGA
145	Contig_1563733_3	954	TAAGCCGTTGTGAGATGCAG	CTGTAACGGGCTGTCCTAGC
146	Contig_1575446	798	TGCTGCACATCTCTTTGACC	GAGAGAGCGTTTCCGTTCTG
147	Contig_2555668	799	TGTCTTGCCGTGTGCATTAC	CCGTCTACCCTGTCGGATAA
148	Contig_92279	842	CCCTAGCACATCCACCTCAT	AGACGCACACAGACGAAATG

Appendix Table 4: Summary of the markers used for GoldenGate Assay.

Marker ID[A]	Close_index[B]	Chromosome	cM	BM4205 population	MM4205 population
ge00001s01	1_0433	1H	93.95	Monomorophic	Polymorphic
ge00002s01	2_0498	2HL	0.00	Monomorophic	Monomorophic
ge00003s01	1_0442	7H	84.92	Monomorophic	Polymorphic
ge00004s01	2_0527	3H	134.31	Monomorophic	Polymorphic
ge00005s01	1_0454	7H	140.21	Monomorophic	Monomorophic
ge00006s01	1_0455	6H	64.36	Monomorophic	Polymorphic
ge00007s01	2_0533	5H	17.38	Polymorphic	Polymorphic
ge00008s01	1_0462	6H	44.77	Monomorophic	Polymorphic
ge00009s01	1_0467	4H	72.08	Monomorophic	Polymorphic
ge00010s01	2_0553	5H	2.81	Monomorophic	Polymorphic
ge00011s01	1_0477	5H	113.11	Monomorophic	Monomorophic
ge00012s01	2_0572	6H	55.65	Polymorphic	Monomorophic
ge00013s01	1_0498	2H	49.07	Monomorophic	Polymorphic
ge00014s01	2_0590	2H	137.51	Polymorphic	Monomorophic
ge00015s01	2_0778	3H	76.20	Polymorphic	Monomorophic
ge00016s01	1_0837	2H	40.50	Monomorophic	Polymorphic
ge00017s01	1_1179	7H	4.12	Monomorophic	Monomorophic
ge00018s01	2_0012	4H	39.76	Monomorophic	Polymorphic
ge00019s01	1_0239	6H	112.32	Monomorophic	Polymorphic
ge00020s01	1_0399	2H	39.10	Monomorophic	Polymorphic
ge00021s01	2_1201	7H	98.50	Monomorophic	Monomorophic
ge00022s01	2_1338	2H	44.84	Polymorphic	Monomorophic
ge00023s01	1_0510	4H	102.37	Monomorophic	Polymorphic
ge00024s01	1_0513	6H	55.94	Polymorphic	Polymorphic
ge00025s01	1_0515	3H	99.89	Monomorophic	Polymorphic
ge00026s01	1_0516	1H	65.53	Polymorphic	Polymorphic
ge00027s01	2_0607	3H	32.83	Polymorphic	Polymorphic
ge00028s01	2_0609	2HS	0.00	Monomorophic	Monomorophic
ge00029s01	2_0595	3H	12.46	Monomorophic	Polymorphic
ge00030s01	1_0525	2H	38.03	Monomorophic	Polymorphic
ge00031s01	2_0612	3H	131.59	Polymorphic	Monomorophic
ge00032s01	1_0539	6H	46.11	Monomorophic	Polymorphic
ge00033s01	2_0625	1H	106.60	Monomorophic	Monomorophic
ge00034s01	2_0628	3H	98.49	Monomorophic	Polymorphic
ge00035s01	2_0629	5H	122.38	Monomorophic	Monomorophic
ge00036s01	1_0559	3H	24.99	Monomorophic	Polymorphic
ge00037s01	2_0644	5HL	0.00	Polymorphic	Polymorphic
ge00038s01	2_0645	5H	87.35	Polymorphic	Polymorphic
ge00039s01	2_0647	3H	43.23	Monomorophic	Monomorophic
ge00040s01	1_0586	1H	121.77	Monomorophic	Polymorphic
ge00041s01	2_0667	2H	72.33	Monomorophic	Polymorphic
ge00042s01	1_0588	4H	89.39	Monomorophic	Polymorphic
ge00043s01	1_0590	1H	138.31	Polymorphic	Monomorophic

ge00044s01	1_0597	1H	42.52	Polymorphic	Monomorophic
ge00045s01	1_0601	3H	46.31	Monomorophic	Polymorphic
ge00046s01	2_0686	5H	172.38	Monomorophic	Monomorophic
ge00047s01	2_0687	6H	124.85	Monomorophic	Polymorphic
ge00048s01	1_0631	3H	144.64	Polymorphic	Polymorphic
ge00049s01	2_0725	6H	105.60	Polymorphic	Monomorophic
ge00050s01	2_0732	4H	92.38	Polymorphic	Monomorophic
ge00051s01	1_0644	1H	127.10	Monomorophic	Polymorphic
ge00052s01	1_0648	2H	41.66	Polymorphic	Monomorophic
ge00053s01	2_0748	2H	56.28	Polymorphic	Polymorphic
ge00054s01	1_0614	4H	100.74	Monomorophic	Polymorphic
ge00055s01	1_0621	5H	37.11	Monomorophic	Polymorphic
ge00056s01	2_0713	5H	65.49	Polymorphic	Monomorophic
ge00057s01	1_0656	2H	128.26	Polymorphic	Monomorophic
ge00058s01	1_0668	4H	44.94	Monomorophic	Monomorophic
ge00059s01	1_0669	6H	2.27	Monomorophic	Monomorophic
ge00060s01	1_0671	5H	59.40	Monomorophic	Polymorphic
ge00061s01	2_0766	5H	46.23	Monomorophic	Monomorophic
ge00062s01	1_0672	3H	37.17	Polymorphic	Monomorophic
ge00063s01	1_0676	6H	28.39	Polymorphic	Polymorphic
ge00064s01	1_0687	7H	140.99	Polymorphic	Monomorophic
ge00065s01	2_0794	3H	26.90	Polymorphic	Monomorophic
ge00066s01	2_0795	5H	108.63	Polymorphic	Polymorphic
ge00067s01	2_0815	4H	76.03	Monomorophic	Monomorophic
ge00068s01	2_0820	4H	63.56	Monomorophic	Monomorophic
ge00069s01	2_0844	1H	108.31	Monomorophic	Monomorophic
ge00070s01	2_0845	5H	39.97	Monomorophic	Polymorphic
ge00071s01	1_0724	4H	82.42	Polymorphic	Monomorophic
ge00072s01	2_0847	7H	140.21	Polymorphic	Monomorophic
ge00073s01	2_0850	5H	102.06	Monomorophic	Polymorphic
ge00074s01	2_0856	3H	56.40	Monomorophic	Monomorophic
ge00075s01	1_0728	3H	62.99	Monomorophic	Polymorphic
ge00076s01	2_0864	2H	31.72	Monomorophic	Monomorophic
ge00077s01	2_0873	5H	26.28	Monomorophic	Polymorphic
ge00078s01	1_0736	5H	180.71	Monomorophic	Polymorphic
ge00079s01	2_0884	5H	137.16	Monomorophic	Monomorophic
ge00080s01	2_0889	6H	75.21	Polymorphic	Polymorphic
ge00081s01	1_0744	1H	26.11	Polymorphic	Polymorphic
ge00082s01	2_0897	5H	182.88	Polymorphic	Polymorphic
ge00083s01	1_0748	6H	123.84	Monomorophic	Polymorphic
ge00084s01	1_0751	4HL	0.00	Polymorphic	Polymorphic
ge00085s01	1_0753	3H	114.00	Monomorophic	Polymorphic
ge00086s01	2_0906	4H	65.05	Polymorphic	Monomorophic
ge00087s01	1_0755	5H	142.20	Monomorophic	Polymorphic
ge00088s01	1_0756	4H	48.50	Polymorphic	Monomorophic

ge00089s01	2_0920	3H	140.91	Monomorophic	Monomorophic
ge00090s01	2_0921	1H	105.85	Monomorophic	Monomorophic
ge00091s01	1_0764	1H	40.99	Polymorphic	Monomorophic
ge00092s01	2_0943	2H	149.61	Polymorphic	Polymorphic
ge00093s01	1_0775	1H	17.26	Polymorphic	Monomorophic
ge00094s01	1_0780	2H	119.05	Monomorophic	Monomorophic
ge00095s01	2_0962	7H	149.80	Polymorphic	Monomorophic
ge00096s01	1_0786	2H	82.75	Monomorophic	Polymorphic
ge00097s01	2_0974	4H	106.03	Monomorophic	Monomorophic
ge00098s01	2_0988	5H	161.58	Polymorphic	Monomorophic
ge00099s01	2_0996	6H	93.12	Monomorophic	Polymorphic
ge00100s01	2_0997	1H	60.19	Monomorophic	Monomorophic
ge00101s01	2_1005	2H	50.49	Monomorophic	Polymorphic
ge00102s01	2_1007	2H	96.25	Polymorphic	Polymorphic
ge00103s01	1_0805	5H	130.13	Monomorophic	Polymorphic
ge00104s01	2_1025	6H	89.57	Polymorphic	Monomorophic
ge00105s01	2_1027	3H	8.86	Polymorphic	Polymorphic
ge00106s01	1_0814	1H	35.45	Polymorphic	Polymorphic
ge00107s01	2_1032	6H	9.06	Monomorophic	Monomorophic
ge00108s01	2_1037	2H	88.74	Monomorophic	Polymorphic
ge00109s01	1_0819	5H	143.92	Monomorophic	Polymorphic
ge00110s01	2_1061	5H	110.26	Monomorophic	Polymorphic
ge00111s01	2_1073	4H	48.50	Polymorphic	Monomorophic
ge00112s01	1_0833	1H	52.46	Monomorophic	Monomorophic
ge00113s01	1_0845	5H	142.20	Monomorophic	Polymorphic
ge00114s01	1_0851	7H	14.96	Monomorophic	Polymorphic
ge00115s01	2_1121	5H	68.35	Monomorophic	Polymorphic
ge00116s01	2_1122	4H	33.38	Monomorophic	Polymorphic
ge00117s01	1_0861	7H	133.79	Monomorophic	Polymorphic
ge00118s01	1_0863	3H	41.68	Monomorophic	Monomorophic
ge00119s01	1_0868	6H	24.36	Monomorophic	Polymorphic
ge00120s01	2_1140	1H	126.01	Monomorophic	Polymorphic
ge00121s01	2_1141	5H	177.07	Polymorphic	Polymorphic
ge00122s01	2_1144	2H	69.25	Monomorophic	Polymorphic
ge00123s01	1_0882	6H	42.36	Polymorphic	Monomorophic
ge00124s01	1_0885	7H	139.72	Polymorphic	Monomorophic
ge00125s01	2_1187	2H	29.15	Polymorphic	Monomorophic
ge00126s01	2_1197	3H	51.73	Monomorophic	Polymorphic
ge00127s01	1_0891	2H	31.02	Monomorophic	Polymorphic
ge00128s01	1_0876	2H	100.37	Polymorphic	Polymorphic
ge00129s01	2_1150	5H	94.43	Monomorophic	Polymorphic
ge00130s01	2_1166	2H	66.83	Monomorophic	Polymorphic
ge00131s01	1_0881	4H	54.25	Monomorophic	Polymorphic
ge00132s01	2_1204	6H	6.07	Monomorophic	Polymorphic
ge00133s01	2_1205	2H	71.56	Monomorophic	Monomorophic

ge00134s01	2_1209	7H	129.91	Monomorophic	Polymorphic
ge00135s01	2_1210	4H	117.60	Monomorophic	Monomorophic
ge00136s01	2_1221	5H	9.33	Polymorphic	Polymorphic
ge00137s01	1_0901	5H	158.37	Polymorphic	Monomorophic
ge00138s01	2_1226	1H	8.77	Polymorphic	Monomorophic
ge00139s01	2_1229	7H	128.36	Monomorophic	Polymorphic
ge00140s01	2_1242	2H	82.75	Monomorophic	Polymorphic
ge00141s01	1_0916	2H	117.91	Monomorophic	Polymorphic
ge00142s01	1_0918	3H	123.68	Monomorophic	Monomorophic
ge00143s01	1_0926	3H	56.40	Monomorophic	Monomorophic
ge00144s01	2_1267	3H	168.40	Monomorophic	Polymorphic
ge00145s01	2_1270	7H	68.46	Monomorophic	Polymorphic
ge00146s01	2_1272	3H	150.37	Monomorophic	Polymorphic
ge00147s01	1_0939	6H	33.74	Monomorophic	Polymorphic
ge00148s01	2_1305	3H	67.57	Polymorphic	Monomorophic
ge00149s01	2_1308	5H	50.27	Polymorphic	Monomorophic
ge00150s01	1_0962	6H	54.60	Polymorphic	Monomorophic
ge00151s01	2_1315	2H	121.50	Polymorphic	Polymorphic
ge00152s01	1_0965	7H	29.82	Monomorophic	Polymorphic
ge00153s01	2_1326	7H	49.68	Monomorophic	Polymorphic
ge00154s01	1_0943	2H	18.32	Monomorophic	Monomorophic
ge00155s01	1_0949	7HS	0.00	Polymorphic	Polymorphic
ge00156s01	2_1339	6H	58.55	Polymorphic	Monomorophic
ge00157s01	1_0987	2H	31.02	Monomorophic	Monomorophic
ge00158s01	2_1353	4H	77.31	Monomorophic	Monomorophic
ge00159s01	1_0988	2H	113.48	Polymorphic	Monomorophic
ge00160s01	2_1354	1H	0.77	Monomorophic	Monomorophic
ge00161s01	2_1363	7H	144.45	Polymorphic	Polymorphic
ge00162s01	1_1227	2H	133.22	Monomorophic	Monomorophic
ge00164s01	1_1243	7H	122.07	Monomorophic	Monomorophic
ge00165s01	1_0994	6H	31.73	Monomorophic	Polymorphic
ge00166s01	2_1373	1H	95.42	Monomorophic	Polymorphic
ge00167s01	2_1374	4H	28.40	Polymorphic	Polymorphic
ge00168s01	2_1377	2H	8.57	Monomorophic	Polymorphic
ge00169s01	2_1384	1H	135.56	Monomorophic	Monomorophic
ge00170s01	2_1385	4H	23.10	Monomorophic	Monomorophic
ge00171s01	2_1388	2H	54.95	Polymorphic	Monomorophic
ge00172s01	2_1399	2H	63.53	Monomorophic	Polymorphic
ge00173s01	2_1401	5H	48.83	Polymorphic	Monomorophic
ge00174s01	1_1014	7H	60.69	Monomorophic	Polymorphic
ge00175s01	1_1015	2H	55.67	Polymorphic	Monomorophic
ge00176s01	2_1419	7HS	0.00	Monomorophic	Monomorophic
ge00177s01	2_1426	5H	27.00	Monomorophic	Polymorphic
ge00178s01	2_1427	3H	141.54	Monomorophic	Monomorophic
ge00179s01	2_1431	1H	64.91	Monomorophic	Polymorphic

ge00180s01	2_1437	7H	17.20	Monomorophic	Polymorphic
ge00181s01	2_1446	1H	92.04	Monomorophic	Polymorphic
ge00182s01	2_1459	2H	127.06	Monomorophic	Monomorophic
ge00183s01	1_1054	2H	50.60	Monomorophic	Monomorophic
ge00184s01	1_1189	1H	90.97	Monomorophic	Monomorophic
ge00185s01	1_1191	3H	64.19	Polymorphic	Monomorophic
ge00186s01	1_1200	5H	117.47	Monomorophic	Monomorophic
ge00187s01	1_1207	4H	62.83	Monomorophic	Polymorphic
ge00188s01	1_1246	6H	81.88	Monomorophic	Monomorophic
ge00189s01	1_1261	6H	65.03	Monomorophic	Monomorophic
ge00190s01	1_1273	5H	111.68	Monomorophic	Polymorphic
ge00191s01	1_1277	1H	97.68	Monomorophic	Polymorphic
ge00192s01	1_1281	5H	63.31	Monomorophic	Monomorophic
ge00193s01	1_1216	5H	171.66	Monomorophic	Polymorphic
ge00194s01	1_1221	5H	60.74	Monomorophic	Polymorphic
ge00195s01	1_1292	4H	96.59	Monomorophic	Monomorophic
ge00196s01	1_1299	4H	111.68	Monomorophic	Polymorphic
ge00197s01	1_1302	2H	52.47	Polymorphic	Polymorphic
ge00198s01	1_1314	3H	70.23	Polymorphic	Polymorphic
ge00199s01	1_1337	3H	56.40	Monomorophic	Monomorophic
ge00200s01	1_1341	5H	113.83	Monomorophic	Monomorophic
ge00201s01	1_1345	4H	5.55	Polymorphic	Polymorphic
ge00203s01	1_1355	5H	86.63	Polymorphic	Polymorphic
ge00204s01	1_1367	1H	66.70	Monomorophic	Polymorphic
ge00205s01	1_1375	5H	130.84	Monomorophic	Monomorophic
ge00206s01	1_1398	4H	87.49	Monomorophic	Monomorophic
ge00207s01	1_1401	3H	58.01	Monomorophic	Monomorophic
ge00208s01	2_1516	7HS	0.00	Monomorophic	Polymorphic
ge00209s01	1_1432	5H	47.39	Monomorophic	Monomorophic
ge00210s01	1_1436	3H	155.85	Monomorophic	Polymorphic
ge00211s01	1_1059	2H	7.14	Monomorophic	Monomorophic
ge00212s01	1_1066	4H	113.92	Monomorophic	Monomorophic
ge00213s01	1_1092	5H	145.35	Monomorophic	Polymorphic
ge00214s01	1_1094	2H	108.61	Monomorophic	Monomorophic
ge00215s01	1_1098	7H	68.46	Monomorophic	Monomorophic
ge00216s01	1_1100	2H	83.82	Polymorphic	Polymorphic
ge00217s01	1_1127	3H	136.66	Monomorophic	Monomorophic
ge00218s01				Polymorphic	Polymorphic
ge00219s01	1_1141	3H	130.82	Monomorophic	Monomorophic
ge00220s01	1_1159	5H	60.74	Monomorophic	Monomorophic
ge00221s01	1_1180	4H	40.96	Monomorophic	Monomorophic
ge00222s01	1_1186	4H	121.83	Monomorophic	Monomorophic
ge00223s01	1_1187	6H	121.22	Monomorophic	Monomorophic
ge00224s01	1_1440	7H	144.45	Monomorophic	Polymorphic
ge00225s01	1_1441	5H	151.36	Monomorophic	Monomorophic

ge00226s01	1_1458	6H	81.17	Polymorphic	Monomorophic
ge00227s01	2_1523	3H	164.29	Polymorphic	Monomorophic
ge00228s01	1_1479	6H	12.54	Monomorophic	Polymorphic
ge00229s01	1_1490	5H	153.60	Polymorphic	Polymorphic
ge00230s01	1_1497	5H	155.13	Monomorophic	Polymorphic
ge00231s01	1_1500	4H	79.58	Polymorphic	Polymorphic
ge00232s01	1_1503	3H	114.00	Monomorophic	Monomorophic
ge00233s01	1_1507	5H	125.81	Monomorophic	Monomorophic
ge00234s01	1_1516	3H	169.32	Monomorophic	Monomorophic
ge00235s01	1_1533	2H	87.33	Polymorphic	Monomorophic
ge00236s01	1_0006	1H	73.94	Polymorphic	Monomorophic
ge00237s01	2_0010	5H	18.72	Polymorphic	Polymorphic
ge00238s01	2_0013	4H	123.29	Monomorophic	Monomorophic
ge00239s01	1_0010	4H	66.00	Polymorphic	Polymorphic
ge00240s01	2_0007	4H	119.09	Monomorophic	Polymorphic
ge00241s01	2_0022	5H	181.43	Polymorphic	Monomorophic
ge00242s01	1_0023	6H	22.35	Polymorphic	Polymorphic
ge00243s01	1_0024	5H	107.59	Monomorophic	Monomorophic
ge00244s01	2_0036	6H	105.60	Monomorophic	Polymorphic
ge00245s01	2_0063	3H	85.99	Monomorophic	Monomorophic
ge00246s01	2_0064	2H	112.91	Monomorophic	Monomorophic
ge00247s01	1_0040	6H	65.03	Monomorophic	Polymorphic
ge00248s01	2_0080	2H	95.64	Polymorphic	Monomorophic
ge00249s01	2_0092	7H	110.99	Polymorphic	Monomorophic
ge00250s01	2_0093	3H	81.66	Monomorophic	Polymorphic
ge00251s01	2_0095	1H	60.77	Monomorophic	Polymorphic
ge00252s01	1_0056	7H	40.18	Monomorophic	Polymorphic
ge00253s01	2_0113	7H	56.81	Monomorophic	Polymorphic
ge00254s01	2_0118	6H	97.39	Monomorophic	Monomorophic
ge00255s01	2_0119	4H	99.28	Monomorophic	Monomorophic
ge00256s01	2_0127	5H	123.52	Monomorophic	Polymorphic
ge00257s01	1_0075	1H	55.49	Monomorophic	Polymorphic
ge00258s01	1_0081	3H	39.45	Polymorphic	Monomorophic
ge00259s01	2_0145	4H	1.64	Monomorophic	Monomorophic
ge00260s01	1_0085	2H	156.72	Polymorphic	Monomorophic
ge00261s01	1_1312	6H	55.00	Polymorphic	Monomorophic
ge00262s01	1_1350	5H	104.50	Monomorophic	Polymorphic
ge00263s01	1_1391	3H	65.52	Polymorphic	Monomorophic
ge00264s01	1_1435	2H	78.03	Polymorphic	Polymorphic
ge00265s01	2_0009	3H	107.63	Polymorphic	Monomorophic
ge00266s01	1_0030	1H	18.05	Monomorophic	Monomorophic
ge00267s01	2_0089	4H	123.29	Monomorophic	Monomorophic
ge00268s01	1_1219	7H	79.60	Monomorophic	Monomorophic
ge00269s01	1_0061	6H	42.36	Monomorophic	Polymorphic
ge00270s01	1_0072	2H	151.37	Polymorphic	Monomorophic

ge00271s01	1_0093	4H	48.50	Polymorphic	Monomorophic
ge00272s01	1_0129	6H	42.36	Monomorophic	Monomorophic
ge00273s01	1_0153	7H	73.75	Monomorophic	Polymorphic
ge00274s01	1_0169	7H	104.78	Monomorophic	Polymorphic
ge00275s01	1_0186	1H	23.86	Polymorphic	Polymorphic
ge00276s01	1_0247	4HL	0.00	Polymorphic	Polymorphic
ge00277s01	1_0280	3H	130.19	Monomorophic	Monomorophic
ge00278s01	2_0367	5H	75.40	Monomorophic	Polymorphic
ge00279s01	1_0327	7H	42.60	Monomorophic	Polymorphic
ge00280s01	2_0392	5H	70.48	Polymorphic	Polymorphic
ge00281s01	2_0480	2H	126.03	Monomorophic	Polymorphic
ge00282s01	2_0162	7H	31.75	Monomorophic	Monomorophic
ge00283s01	2_0169	1H	99.95	Monomorophic	Monomorophic
ge00284s01	2_0172	3H	16.33	Polymorphic	Polymorphic
ge00285s01	1_0107	6H	119.02	Monomorophic	Monomorophic
ge00286s01	1_0109	2H	127.64	Monomorophic	Monomorophic
ge00287s01	2_0192	7H	34.82	Monomorophic	Polymorphic
ge00288s01	2_0159	3H	2.90	Polymorphic	Polymorphic
ge00289s01	1_0095	5H	137.16	Monomorophic	Polymorphic
ge00290s01	1_0113	4H	19.52	Polymorphic	Monomorophic
ge00291s01	2_0193	3H	42.06	Monomorophic	Polymorphic
ge00292s01	2_0229	1H	71.43	Polymorphic	Monomorophic
ge00293s01	1_0143	7H	87.97	Monomorophic	Monomorophic
ge00294s01	2_0232	6HS	0.00	Polymorphic	Monomorophic
ge00295s01	2_0236	5H	80.61	Monomorophic	Polymorphic
ge00296s01	1_0165	6H	16.97	Monomorophic	Monomorophic
ge00297s01	2_0245	7H	12.42	Monomorophic	Polymorphic
ge00298s01	2_0247	7H	116.33	Monomorophic	Polymorphic
ge00299s01	2_0249	7H	52.82	Monomorophic	Monomorophic
ge00300s01	2_0252	3H	6.03	Polymorphic	Monomorophic
ge00301s01	1_0174	7H	166.56	Monomorophic	Polymorphic
ge00302s01	1_0175	6H	119.02	Monomorophic	Polymorphic
ge00303s01	1_0180	2H	21.61	Monomorophic	Polymorphic
ge00304s01	2_0266	6H	59.56	Polymorphic	Monomorophic
ge00305s01	1_0202	6H	90.15	Monomorophic	Monomorophic
ge00306s01	1_0214	2H	93.50	Monomorophic	Polymorphic
ge00307s01	1_0438	1H	50.00	Monomorophic	Monomorophic
ge00308s01	2_0531	6H	97.39	Monomorophic	Polymorphic
ge00309s01	2_0558	6H	112.32	Polymorphic	Polymorphic
ge00310s01	1_0523	4H	78.77	Polymorphic	Polymorphic
ge00311s01	2_0620	6H	70.04	Monomorophic	Polymorphic
ge00312s01	2_0493	6H	1.34	Polymorphic	Monomorophic
ge00313s01	1_0565	3H	19.15	Polymorphic	Monomorophic
ge00314s01	2_0659	3H	91.25	Monomorophic	Polymorphic
ge00315s01	2_0676	5H	146.00	Monomorophic	Polymorphic

ge00316s01	1_0606	4H	67.46	Polymorphic	Monomorophic
ge00317s01	2_0755	7H	15.93	Monomorophic	Polymorphic
ge00318s01	2_0771	7H	87.21	Monomorophic	Polymorphic
ge00319s01	1_0688	5H	34.25	Monomorophic	Polymorphic
ge00320s01	2_0829	5H	168.79	Polymorphic	Monomorophic
ge00321s01	1_0738	4H	19.52	Polymorphic	Monomorophic
ge00322s01	1_0747	3H	93.43	Monomorophic	Monomorophic
ge00323s01	2_0908	1H	121.12	Monomorophic	Polymorphic
ge00324s01	1_0760	1H	34.83	Monomorophic	Polymorphic
ge00325s01	1_0782	1H	131.89	Monomorophic	Polymorphic
ge00326s01	2_1014	6H	54.60	Polymorphic	Monomorophic
ge00327s01	1_0817	6H	45.44	Monomorophic	Polymorphic
ge00328s01	2_0293	2H	147.94	Monomorophic	Polymorphic
ge00329s01	1_0221	4H	21.61	Monomorophic	Polymorphic
ge00330s01	2_0298	5H	132.63	Monomorophic	Polymorphic
ge00331s01	2_0306	5H	58.70	Monomorophic	Polymorphic
ge00332s01	1_0244	6H	43.15	Monomorophic	Monomorophic
ge00333s01	2_0320	5H	108.18	Polymorphic	Monomorophic
ge00334s01	1_0262	4H	55.63	Monomorophic	Monomorophic
ge00335s01	1_0265	2H	70.54	Polymorphic	Polymorphic
ge00336s01	1_0270	6H	60.23	Polymorphic	Polymorphic
ge00337s01	1_0287	2H	85.92	Monomorophic	Polymorphic
ge00338s01	2_0355	6H	110.32	Polymorphic	Polymorphic
ge00339s01	1_0292	5H	144.63	Monomorophic	Polymorphic
ge00340s01	1_0303	7H	87.97	Monomorophic	Polymorphic
ge00341s01	1_0312	3H	114.00	Monomorophic	Polymorphic
ge00342s01	1_0315	2H	141.28	Monomorophic	Polymorphic
ge00343s01	1_0319	4H	8.25	Monomorophic	Monomorophic
ge00344s01	2_0379	6H	101.44	Monomorophic	Monomorophic
ge00345s01	1_0331	6H	81.22	Monomorophic	Monomorophic
ge00346s01	1_0332	1H	15.40	Polymorphic	Polymorphic
ge00347s01	1_0335	3H	65.52	Polymorphic	Monomorophic
ge00348s01	1_0336	5H	161.58	Monomorophic	Polymorphic
ge00349s01	2_0394	2H	27.29	Polymorphic	Monomorophic
ge00350s01	1_0342	2H	44.13	Polymorphic	Monomorophic
ge00351s01	2_0415	6H	13.21	Monomorophic	Polymorphic
ge00352s01	1_0383	2H	130.01	Monomorophic	Monomorophic
ge00353s01	2_1077	5H	146.00	Monomorophic	Polymorphic
ge00354s01	1_0854	1H	117.80	Monomorophic	Polymorphic
ge00355s01	2_1133	5H	80.61	Monomorophic	Monomorophic
ge00356s01	2_1145	3H	42.47	Monomorophic	Monomorophic
ge00357s01	2_1203	5H	129.41	Monomorophic	Polymorphic
ge00358s01	2_1212	3H	111.42	Monomorophic	Polymorphic
ge00359s01	1_0909	2H	63.53	Polymorphic	Monomorophic
ge00360s01	2_1261	2H	28.44	Polymorphic	Monomorophic

ge00361s01	2_1330	7H	86.44	Monomorophic	Polymorphic
ge00362s01	2_1350	5H	51.00	Polymorphic	Monomorophic
ge00363s01	2_1355	5H	153.51	Polymorphic	Monomorophic
ge00364s01	1_0999	7H	161.43	Polymorphic	Monomorophic
ge00365s01	2_1392	1H	114.84	Polymorphic	Monomorophic
ge00366s01	1_1048	5H	29.90	Polymorphic	Monomorophic
ge00367s01	1_1198	5H	48.83	Polymorphic	Monomorophic
ge00368s01	1_1067	6H	58.01	Polymorphic	Monomorophic
ge00369s01	1_1111	6H	128.48	Monomorophic	Monomorophic
ge00370s01	1_1172	3H	126.27	Polymorphic	Monomorophic
ge00371s01	1_0578	5H	95.08	Monomorophic	Polymorphic
ge00372s01	1_0791	2H	150.67	Polymorphic	Monomorophic
ge00373s01	2_0267	1H	101.45	Monomorophic	Polymorphic
ge00374s01	2_0449	5H	100.28	Monomorophic	Monomorophic
ge00375s01	2_0455	3H	28.44	Monomorophic	Monomorophic
ge00376s01	1_0404	2H	117.20	Monomorophic	Polymorphic
ge00377s01	1_0409	4H	3.74	Polymorphic	Polymorphic
ge00378s01	1_0419	1H	3.75	Monomorophic	Polymorphic
ge00379s01	1_0427	6H	34.40	Polymorphic	Monomorophic
ge00380s01	1_0387	4H	119.84	Monomorophic	Monomorophic
ge00381s01	2_0434	1H	88.23	Monomorophic	Polymorphic
ge00382s01	1_0396	1H	96.92	Monomorophic	Polymorphic
ge00383s01	1_0429	2H	115.08	Polymorphic	Monomorophic
ge00384s01	2_0495	7H	25.70	Monomorophic	Polymorphic

[A]- Marker name used in the Illumina GoldenGate Assay.

[B]- Corresponding marker name according to Close et al. (2009).

Appendix Table 5: List of the sequenced MTP BAC clones.

In-house ID	Published ID	Chromosome	CB_start	CB_end	FP_Contig
MNRA0395M21	HVVMRXALLrA0395M21	7	0	146	7112
MNHA0090J05	HVVMRXALLhB0090J05	7	31	123	7112
MNRA0043E12	HVVMRXALLrA0043E12	7	52	213	7112
MNRA0178K21	HVVMRXALLrA0178K21	7	73	116	7112
MNRA0180I18	HVVMRXALLrA0180I18	7	76	113	7112
MNMA0223J05	HVVMRXALLmA0223J05	7	166	279	7112
MNMA0037P04	HVVMRXALLmA0037P04	7	213	344	7112
MNEA0017I24	HVVMRXALLeA0017I24	7	273	373	7112
MNMA0159P16	HVVMRXALLmA0159P16	7	286	391	7112
MNEA0098N12	HVVMRXALLeA0098N12	7	359	426	7112
MOHI0023E08	HVVMRXALLhA0023E08	7	393	486	7112
MNRA0254F14	HVVMRXALLrA0254F14	7	0	134	7615
MNEA0172A05	HVVMRXALLeA0172A05	7	53	134	7615
MNMA0220E09	HVVMRXALLmA0220E09	7	22	123	7615
MNMA0230A06	HVVMRXALLmA0230A06	7	62	159	7615
MOHI0032I06	HVVMRXALLhA0032I06	7	63	169	7615
MNHA0145H15	HVVMRXALLhB0145H15	7	0	99	4483
MNHA0138E15	HVVMRXALLhB0138E15	7	45	142	4483
MNRA0045N01	HVVMRXALLrA0045N01	7	78	156	4483
MNMA0337G17	HVVMRXALLmA0337G17	7	109	211	4483
MNMA0155C21	HVVMRXALLmA0155C21	7	174	289	4483
MNMA0299O06	HVVMRXALLmA0299O06	7	258	382	4483
MNMA0465P05	HVVMRXALLmA0465P05	7	312	443	4483
MNMA0510H04	HVVMRXALLmA0510H04	7	378	508	4483
MNMA0381J10	HVVMRXALLmA0381J10	7	425	556	4483
MNEA0143C22	HVVMRXALLeA0143C22	7	496	590	4483
MNRA0400I23	HVVMRXALLrA0400I23	7	0	133	44845
MNRA0226A20	HVVMRXALLrA0226A20	7	27	114	44845
MNEA0131L13	HVVMRXALLeA0131L13	7	77	157	44845
MNHB0124B19	HVVMRXALLhC0124B19	7	128	211	44845
MNEA0089H17	HVVMRXALLeA0089H17	7	130	235	44845
MNMA0510A15	HVVMRXALLmA0510A15	7	194	285	44845
MNMA0037G09	HVVMRXALLmA0037G09	7	223	351	44845
MNMA0169C16	HVVMRXALLmA0169C16	7	225	303	44845
MNMA0089A11	HVVMRXALLmA0089A11	7	339	442	44845
MNHA0142P06	HVVMRXALLhB0142P06	7	398	471	44845
MNMA0272A07	HVVMRXALLmA0272A07	7	432	532	44845
MNMA0180L03	HVVMRXALLmA0180L03	7	489	601	44845
MNMA0416P19	HVVMRXALLmA0416P19	7	554	692	44845
MOKI0095B11	HVVMRX83KhA0095B11	7	621	723	44845
MNMA0457P05	HVVMRXALLmA0457P05	7	679	776	44845
MNHA0062D15	HVVMRXALLhB0062D15	7	712	794	44845
MNMA0268L16	HVVMRXALLmA0268L16	7	744	865	44845

MNEA0192I19	HVVMRXALLeA0192I19	7	810	933	44845
MNHA0069D14	HVVMRXALLhB0069D14	7	849	1048	44845
MNHA0103F17	HVVMRXALLhB0103F17	7	0	114	7506
MOHI0054F13	HVVMRXALLhA0054F13	7	1	56	7506
MNMA0125F11	HVVMRXALLmA0125F11	7	66	171	7506
MNEA0272A06	HVVMRXALLeA0272A06	7	79	172	7506
MNMA0208I02	HVVMRXALLmA0208I02	7	124	226	7506
MOHI0037N16	HVVMRXALLhA0037N16	7	163	227	7506
MNMA0258A21	HVVMRXALLmA0258A21	7	197	264	7506
MNEA0213D22	HVVMRXALLeA0213D22	7	223	328	7506
MNMA0340A18	HVVMRXALLmA0340A18	7	251	365	7506
MNEA0135I17	HVVMRXALLeA0135I17	7	307	417	7506
MNMA0477O22	HVVMRXALLmA0477O22	7	352	463	7506
MNEA0275D08	HVVMRXALLeA0275D08	7	415	522	7506
MNRA0380O16	HVVMRXALLrA0380O16	7	438	469	7506
MNEA0188J13	HVVMRXALLeA0188J13	7	464	539	7506
MNMA0412B01	HVVMRXALLmA0412B01	7	465	566	7506

8 Abbreviations

%	Percent
˚C	Celsius
µg	Microgram
µl	Microliter
µM	Micromolar
AA	Amino Acids
ATP	Adenosine triphosphate
BAC	Bacterial Artificial Chromosome
BLAST	Basic Local Alignment Search Tool
BOPA	Barley Oligonucleotide Pool Assay
bp	base pairs
CAPS	Cleaved Amplified Polymorphic Sequences
CB	Consensus Band
cDNA	Complementary DNA
ChIP-seq	Chromatin immunoprecipitation sequencing
cm	Centimeter
cM	Centimorgan
CTAB	Cetyltrimethylammonium bromide
cTP	Chloroplast Transit Peptide
cv.	Cultivar
DEPC-H_2O	Diethylpyrocarbonate-treated water
DH	Doubled Haploid
DNA	Deoxyribonucleic acid
dRNA-seq	differential RNA sequencing
DSB	Double Strand Break
dsRNA	double-stranded RNA
EDTA	Ethylenediaminetetraacetic acid
EMS	Ethyl Methane Sulfonate
EST	Expressed Sequence Tags
F_1	First generation
F_2	Second filial generation
F_3	Third filial generation
FLcDNA	Full-length cDNA

FPKM	Fragments per kilobase of exon per million fragments mapped
g	Gram
GBS	Genotyping By Sequencing
GFP	Green Fluorescence Protein
GWAS	Genome-wide Association Study
IBSC	International Barley Genome Sequencing Consortium
ID	Identifier
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research
IR	Inverted Repeat
Kbp	Kilobase pairs
LOD	Logarithm of Odds
LSC	Large Single Copy
M	Molar
M_2	Second generation of EMS-treated seeds
M_3	Third generation of EMS-treated seeds
M_4	Fourth generation of EMS-treated seeds
Mbp	Megabase pairs
mg	Milligram
min	Minute(s)
ML	Maximum Likelihood
ml	Milliliter
mM	Millimolar
MOPS	3-(N-morpholino) propanesulfonic acid
mRNA	Messenger RNA
MT	Mutant Type
MTP	Minimal Tiling Path
mTP	Mitochondrial Transit Peptide
NCBI	National Center for Biotechnology Information
ndh	Nicotinamide dehydrogenase
NEP	Nucleus-encoded Plastid RNA Polymerase
NGRC	Nordic Genetic Resource Center
NGS	Next Generation Sequencing
nm	Nanometer
P	Parents
p	Possibility

PCR	Polymerase Chain Reaction
PEP	Plastid-encoded Plastid RNA Polymerase
POPSEQ	Population Sequencing
RIL	Recombination Inbred Line
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RuBisCO	Ribulose-1,5-bisphosphate carboxylase
SNP	Single Nucleotide Polymorphism
SP	Secretory Pathway
SSC	Small Single Copy
TALEN	Transcription Activator-like Effector Nuclease
T-DNA	Transfer DNA
TE	Transposable Element
TEM	Transmission Electron Microscopy
TILLING	Targeting Induced Local Lesions in Genomes
Tm	DNA melting temperature
tRNA	Transfer RNA
U	Unit
V	Voltage
v/v	volume/volume
w/v	weight/volume
WGS	Whole Genome Shotgun
WT	Wild-type
Y2H	Yeast two-hybrid

9 Acknowledgments

First of all, I do appreciate the China Scholarship Council (CSC) for its financial support, with which I could have the privileged opportunity to pursue my PhD studies abroad. The four-year experience in Germany would definitely be one of the most memorable moments in my life.

My PhD work was performed in the research group of Genome Diversity (GED) at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben. I would like to express my sincere appreciation to my supervisor Dr. Nils Stein for his patient guidance. His strong knowledge, great scientific enthusiasm, and cautious attitude have supported my flourishing self-development.

Exceedingly, I would like to thank my supervisor Prof. Dr. Thomas Börner, Humboldt Universität zu Berlin, for his terrific interest in my work, which made the completion of my dissertation possible. I am grateful to him for the profound discussion, kind consideration, and general support of my PhD project.

I owe my sincere gratitude to my colleague and friend Matthias Jost, for all kinds of support and especially his generous help at early stage of my PhD to get used to the working environment.

I do feel lucky to have these excellent people being my colleagues, Neele Wendler, Dr. Ruonan Zhou, Dr. Axel Himmelbach, Dr. Ping Yang, and Dr. Ruvini Ariyadasa. The valuable support and suggestions from your side are kindly acknowledged.

The excellent technical assistance from Mary Ziems, Jacqueline Pohl, Jelena Perovic, Manuela Knauft, Sandra Driesslein, Susanne König and Ines Walde had made my project running smoothly.

I also would like to take this opportunity to convey my sincere and wholehearted thanks to all the in-house collaborators. Dr. Jochen Kumlehn, Dr. Götz Hensel, Stefan Hiekel, Hannes Trautwein, and Ingrid Otto from research group Plant Reproductive Biology for their support on the transgenic experiments; Dr. Martin Mascher and Sebastian Beier from research group Bioinformatics and Information Technology for their assistance on NGS data analysis; Dr. Michael Melzer and Dr.

Twan Rutten from research group Structure Cell Biology for their well-organized SEM and TEM courses and image analysis.

In addition, I am especially grateful to Dr. Viktor Korzun, KWS LOCHOW GmbH for performing the Illumina GoldenGate genotyping assay.

My special thanks are given to the gardeners, for their heed attention to my plant material in the greenhouse.

Moreover, it is my pleasure to say ‘Thank you’ to Dr. Britt Leps and Ms. Jacqueline Heidecke for taking care of various aspects of the administrative issues, which build me a convenient living atmosphere.

Besides science, the ‘badminton and basketball team’ do bring me lots of fun to my daily life. Thanks to you all.

Last but not least, I am deeply indebted to my beloved parents for their continuous support, understanding and inspiration.

10 Eidesstattliche Erklärung

Hiermit versichere ich, die vorliegende Dissertation eigenständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben. Die dem Verfahren zugrunde liegende Promotionsordnung ist mir bekannt.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

…… ……..
(Date) (Mingjiu Li)